Understanding the catalytic mechanism of ribokinase from *Leishmania donovani* using structural and functional approach

Thesis submitted to the University of Hyderabad for the degree of

Doctor of Philosophy

By

Santhosh Gatreddi

(Regd No. 11LTPH13)

Supervisor

Dr. Insaf Ahmed Qureshi

Department of Biotechnology and Bioinformatics
School of Life Sciences
University of Hyderabad
Hyderabad-500046, India

December 2018

University of Hyderabad
Hyderabad- 500 046, India

CERTIFICATE

This is to certify that this thesis entitled ***"Understanding the catalytic mechanism of ribokinase from Leishmania donovani using structural and functional approach"*** submitted by **Mr. Santhosh Gatreddi** bearing registration number **11LTPH13** in partial fulfilment of the requirements for award of Doctor of Philosophy in the Department of Biotechnology and Bioinformatics, School of Life Sciences, is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Parts of this thesis have been:

A. Published in the following publications:

1. **Gatreddi S**, Are S, & Qureshi IA. (2018), *Acta Crystallogr F Struct Biol Commun.* Feb1; 99:104. (PMID: 29400319), Chapter 1-3.

B. Presented in the following conferences:

1. 3rd International conference on innovations in pharmaceutical sciences 2018, Aug 3-4, Hyderabad, India.
2. Bioquest 2016, Oct 20-21, School of Life Sciences, University of Hyderabad, Hyderabad, India.

Further, the student has passed the following courses towards fulfilment of coursework requirement for Ph.D.

Course code	Name	Credits	Pass/Fail
BT 801	Seminar 1	1	PASS
BT 802	Research Ethics & Management	2	PASS
BT 803	Biostatistics	2	PASS
BT 804	Analytical Techniques	3	PASS
BT 805	Lab work	4	PASS

Supervisor **Head of Department** **Dean of School**

University of Hyderabad
Hyderabad- 500 046, India

DECLARATION

I, **Santhosh Gatreddi**, hereby declare that this thesis entitled ***"Understanding the catalytic mechanism of ribokinase from Leishmania donovani using structural and functional approach"*** submitted by me under the guidance and supervision of **Dr. Insaf Ahmed Qureshi**, is an original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Date:

Signature of the student

Name: Santhosh Gatreddi

Regd No. 11LTPH13

Signature of the Supervisor

Acknowledgements

Completion of this doctoral dissertation would not have been possible without the support of several people. I would like to express my sincere thanks to all of them.

First of all, I would like to express my deepest gratitude towards my research supervisor, **Dr. Insaf Ahmed Qureshi**, for his constant guidance throughout my Ph.D. His support during the critical phases of my Ph.D is phenomenal. He lightened up the way in my darkest times and encouraged me a lot in the academic life. I am also very grateful for his scientific advice, knowledge and many insightful discussions about research and life. The everlasting discussions and the freedom of my thoughts to design and plan experiments made me more research oriented and gave a lot of self-confidence to build my own research aptitude. He always made himself available to clarify doubts despite his busy schedules, and I consider it is a great opportunity to do my doctoral programme under his guidance and to learn from him research expertise.

I would like to extend my sincere gratitude to our collaborator **Dr. Anthony Addlagatta, CSIR-IICT** for his continuous support, guidance, cooperation and encouragement. His enthusiasm and integral view on research made a deep impression on me. I would like to thank for his valuable suggestions to carry out my third objective successfully and this work would not be possible without his help.

It gives me immense pleasure to thank another collaborator of my work **Dr. Dileep Vasudevan, ILS** for his constant support and encouragement to carry out my structural work. I am grateful for his kind help at all stages of my Ph.D work by providing crystallization accessories, collecting high resolution data sets, sharing insights into the structure analysis and writing. I am grateful for the generous sharing of the research experience and plentiful suggestions on the crystallization.

I am also thankful to **Mr. Vijay Kumar Pillalamarri**, who helped me in the crystallization, data collection and structure determination throughout my work at IICT, Hyderabad.

Besides my mentor, I would like to thank my doctoral committee members, **Dr. N. Prakash Prabhu** and **Dr. Mohd. Akif** for their insightful evaluation and critical suggestions for improving my work.

I am also thankful to **Dr. Ravindra D. Makde, Dr. Ashwani Kumar and Dr. Biplav Ghosh** for their kind help during data collection and processing at RRCAT, Indore.

I am also greatful to **Dr. Irfan Ahmad Ghazi** for allowing me to access the lab facilities, constant encouragement and kind support to carry out my Ph.D work.

I am also thankful to **Drs. Mohd. Akif, Nooruddin Khan** and **Bindu Madhava Reddy for** allowing me to use their lab facilities.

I wish to thank the present and former **Heads and Deans** of the **Department of Biotechnology and Bioinformatics and School of Life Sciences**, University of Hyderabad, respectively for providing me with all the essential facilities.

I am thankful to my lab members, **Sayanna, Pranay, Saleem, Narsimulu, Janish** for their support and for creating a peaceful and amiable work atmosphere in the lab. I am also thankful to present and former M.Sc. students of our lab for their help throughout my work.

I would like to extend my thanks to **Dr. Sudhaker Dharavath, JNU** for his valuable suggestions during my structural work.

I am also thankful to **Mr. Tejaswi Koduru** and for his help during CD experiments, data fitting and for scientific discussions.

I would like to thank my friends **Dr. Rambabu, Bantu, Jagadeesh, Surya, Kiran, Dr. Kishore, Dr. Shanmukh Kumar, Dr. Mohan Raj, Somashekar, Sireesh, Dr. Yasin, Dr. Srikanth, Dr. Sravan, Dr. Suman, Dr. Srinivas, Dr. Trinadh, Just For Fun team and other SLS friends** for their support and making the campus life cheerful and memorable.

I also extend my thanks to Biotechnology office staff, **Mr. Rajashekar, Mr. Shekar, and Mr. Rahul** for their help in time.

I greatly acknowledge UGC for providing me financial support (JRF and SRF).

I also thank all the funding bodies (CSIR, DST, UPE, DBT and PURSE) for their financial assistance to the ldepartment and school.

I am much indebted to my parents and sisters for their love and support.

Santhosh Gatreddi

Dedicated to My Parents

CONTENTS

Page No.

List of Figures and Tables

Figures **Page No.**

Tables

Abbreviations

µl	Micro litre
µM	Micro molar
Å	Angstrom
ADP	Adenosine 5’ diphosphate
AMPPCP	5'-Adenylyl (beta, gamma-methylene)diphosphonate
ATP	Adenosine 5' triphosphate
bp	Base pair
CCP4	Collaborative Computational Project 4
CD	Circular Dichroism
COOT	Crystallographic Object-Oriented Toolkit
DNA	Deoxy ribonucleic acid
EDTA	Ethylene diamine tetraacetic acid
K	Kelvin
kb	Kilo Base
kDa	Kilo Dalton
M	Molar
mg	Milli gram
min	Minutes
ml	Milli litre
mM	Milli molar
NADH	Nicotinamide adenine dinucleotide
nm	Nano metre
OD	Optical density
PDB	Protein data bank
PfkB	Phosphofructokinase B
RK	Ribokinase
SDS-PAGE	Sodium dodecyl sulphate-Polyacrylamide gel electrophoresis

1. Introduction and Review of literature

1. Introduction and review of literature

1.1 History

The protozoan parasites belonging to the order kinetoplastida cause spectrum of diseases to humans [1]. Trypanosomatina is one of the two sub-orders belonging to the kinetoplastida which constitutes several parasites of the nine genera which infects plants, vertebrates and all classes of invertebrates [2]. Out of which the two genera namely, *Trypanosoma* and *Leishmania* causes infections to humans [3]. *Leishmania* was discovered independently by a Scottish pathologist Leishman (1865-1926) and the Irish doctor Donovan (1863-1951) while observing specimen of a patients died due to splenomegaly in India [4, 5]. The parasite was initially identified as ovoid body, which was later confirmed as a novel protozoan by the Ronald Ross [6]. Out of 30 different species of *Leishmania*, about 21 species can cause spectrum of diseases to humans. It is one of the neglected tropical diseases embedded in poverty [7]. The term leishmaniasis portrays a spectrum of diseases caused by the protozoan parasites belonging to the genus *Leishmania*. Leishmaniasis has been grouped into two major categories based on the geographic distribution of infection, Old world and New world. The disease associated with *Leishmania* species found in the Africa, Mediterranean basin and the Middle East refers to the "Old World" leishmaniasis [8]. "New World" leishmaniasis is associated with the species found in Central America, Mexico and South America [9]. Depending on the infecting species, pathogenicity, virulence and clinical manifestations also varies [10].

1.2 Pathogenesis

Various clinical manifestations observed in leishmanaisis is due to the interplay between infecting species type and immune response of the host [11]. The differences in the genomic composition arised during the evolution could be a probable reason for this variability, making a particular species to invade visceral organs and others to target the skin [12]. Though it is poorly understood, the type of macrophages being infected by the visceral and cutaneous *Leishmania* species differ based on the parasite's affinity to the surface receptor of specific macrophage [13, 14]. In addition, clinical manifestations are also affected by the host's competence to respond through adaptive and innate immunity which is largely determined by genetic and acquired factors [11].

1.2.1 Cutaneous Leishmaniasis (CL)

CL is the most prevalent clinical manifestation of leishmaniasis. Based on the Eurocentric world view, the parasites can be divided into Old Word species : *L.infantum, L.major, and L. (L.) tropica* and New World Species, which are *L. (L.) amazonensis, L. (V.) braziliensis, L. (L.) chagasi, L. (V.) guyanensis, L. mexicana L, and L. (V.) naiffi* [15]. Mode of transmission of cutaneous leishmaniasis is either anthroponotic or zoonotic [16]. The subgenus *Leishmania* or *Vianna*, were classified based on the attachment and site development in the sandfly gut [17]. The incubation period for CL in the old world varies from two weeks to several months and may last upto three years as reported in some cases, in new world it is usually 2-8 weeks [18, 19]. CL mostly occurs on the exposed body parts such as face, neck and limbs which are accessible to sandfly bite. Several factors that influence the appearance of symptoms in infected individuals include the type of species, host genetic composition and extent of immune response [11]. The infection caused by the *L. major* and *L. tropica* is less severe and heal quickly. Irrespective of species, CL is not fatal. Direct contact with the skin lesion is harmless, as infection needs inoculation of substance from active sores and lesion are not painful as they look like [10]. It was attempted to classify the several clinical manifestations of CL caused by various *Leishmania* species in the Old and New world groupings, nevertheless due to the overlap these are useful to clinically distinguish species [10]. Incomplete treatment of primary skin lesions may progresses to mucocutaneous leshmaniasis in some cases [10].

1.2.2 Mucocutaneous leishmaniasis (ML)

Mucocuntaneous leishmaniasis (ML) also known as “espundia” is found in Latin America and to a minor extent in Africa (Ethiopia and Sudan) [20]. *Leishmania* species belongs to the *Vianna* subgenus usually found in the Americas, such as *Leishmania* (*Viannia*) *amazonensis, L. (V) braziliensis, Leishmania (V) guyanensis, L. (V) panamensis* are known risk factors for ML [10]. Primary ulcerative lesions (or non ulcerative depends on the *Leishmania* species found out of Americas) appear as either in singles or in multiples [21] and persistent nasal congestion is the most commonly manifested symptom [22].The ML is caused after the months or years of primary cutaneous lesion appearance followed by disease reactivation. ML results in the facial deformations and functional alterations, which reflects the morbidity of disease in social and psychological stigma [23]. It is more lethal compared to CL and fatal if left untreated [7]. ML is

characterized by strong devastating lesions appear on the nasal septum and palate as a result of strong immunopathological response [16]. The parasite spreads to oral and respiratory tract mucosae through lymphatic or heamatogenous route either concurrently or after months to years [11]. Among the population infected with CL, 1-10% of patients can have infection that is extended to the mucosa [19, 21, 24, 25]. Around 90% of ML cases present a scar from the preceding CL incidence [26, 27], which might have happened decades prior [28].

1.2.3 Visceral leishmaniasis (VL)

Because of *L. donovani*, humans became the major reservoirs for the VL. The incubation period for VL is variable and following exposure the clinical manifestations may visible from 10 days to less than 1 year [29]. The common symptoms appear in children of old and new world VL include weight loss, fever, abdominal distension, anorexia and weakness that advance over weeks to months [30, 31]. Distribution of parasites leads to the accumulation of mononuclear phagocytes within the invaded organs and the manifestation of reticulo-endothelial hyperplasia as splenomegaly, hepatomegaly which are the classical signs of VL [32]. The children infected with VL in Sudan exhibits an additional sign called lymphoadenopathy [33]. Regardless of suitable treatment, VL may relapse even after 6-12 months [11]. In the case of untreated VL, the disease may progress to secondary infections, severe wasting, low platelet count and eventually leads to death within 2-3 years [34]. Nutritional status and age are the other properties influencing host response [35]. Severe malnutrition or wasting observed in patients with high parasite load, especially in children [43], but it is not completely understood whether this is effect of poor nutrition or its consequence [44]. To combat the future infection, the parasite should persist both at the site of primary infection and/or draining to lymphnodes after healing [36]. Though it safeguards against reinfection, there might be an underlying risk of relapse reversion if host immunity weakens [10]. VL is also called as kala-azar in Indian subcontinent due to the increased pigmentation of skin as a result of over-production of adrenocorticotropin hormone [37]. *Leishmania* and HIV share similar immunepathological mechanism affecting macrophages and dendritic cells, that results in the advancement of both diseases owing to increased pathogen multiplication [38]. According to WHO guidelines, any form of VL coinfected with HIV should be considered as the atypical disseminated leishmaniasis [39], this condition might present parasites from the liver, respiratory tract and gastrointestinal mucosa [40]. One of the major

challenges to control VL is the co-infection with HIV. In Ethiopia, co-infection of upto 18% were reported [41], while recent reports from Brazil and India have shown that upto 16% of patients presenting with VL are co-infected [42, 43].

1.2.4 Post kala-azar dermal leihsmaniasis (PKDL)

The late complication of VL caused by *L. donovani* is called as post-kala-azar dermal leishmaniasis (PKDL). This condition is infrequent in patients infected with *L. infantum*, until unless people are immunocomprised [44]. PKDL is a consequence of an immunological reaction to the parasites persists in the skin and acts as a reservoir for transmission. In PKDL, parasites seem to sustain in the skin even after treatment where T-helper-2 response observed with elevated levels of interleukin 10, while T-helper-1 response dominates systematically [16]. Infected individuals present with a hypopigmented erythematous maculopapular or macular rash all over the mouth and the trunk that can be extended further to whole body [45]. Less than 10% of VL cases occur in Asia (Bangladesh, India, Nepal) and 60% in east Africa (Ethiopia and Sudan) are complicated further to dermatosis with its varying clinical spectrum between the two regions [46]. The data for treatment regimens for PKDL is rare, and the lack of consistent endpoints makes it difficult to correlate [16]. Generally, PKDL in east Africa is self-curable in majority of patients except for severe cases, conversely, it is treated in Indian subcontinent [16].

1.3 Prevalence

Leishmaniasis was identified as an emerging and uncontrolled category I disease according to the World Health Organization (WHO) [15]. The actual disease burden is not precise due to the lack of diagnostic methods or suitable epidemiological surveillance and the available incidence reports mostly based on the estimates [11]. *Leishmania* infects 12-15 million people throughout the world with a distribution over 89 countries [47, 48]. The continents which are endemic for the disease are Africa, America, Asia and the Mediterranean region [49]. It claims 70,000 lives and infects 1.5-2 million people each year [47, 50]. Only seven countries (Algeria, Afganistan, Brazil, Iran, Syria and Saudi Arabia) contribute to 90% of CL cases throughout the world [51]. The total number of new cases occur annually are 500,000–1,000,000 but only a minor number of cases (19%–37%) are reported to the health authorities [52].

L. donovani is responsible for causing visceral leishmaniasis (VL) in Asia and Africa, whereas *L. infantum* in Central America, South America, the Middle East and Mediterranean Basin[16]. More than 90% of global VL burden was reported from seven countries (India, Brazil, Kenya, Sudan, South Sudan and Somalia), though the disease is prevalent in more than 60 countries [16]. Overwhelming epidemics have been reported from Indian subcontinent [53] and east Africa [54]. The worldwide incidence of VL diminished substantially from 2012 (2,00,000 and 4,00,000 new cases) to 2017(50,000 and 90,000) [52, 55]. India, Bangladesh and Nepal, which contributes more than 50% of global disease burden, have committed to reduce the incidence below one for every 10,000 people per annum at the sub-district (Bangladesh and India) and district level (Nepal) in order to eliminate VL [56]. In endemic areas the disease is asymptomatic; the sero-prevalence for *L. donovani* varies from 7-63% in Indian sub-continent [57], and for *L. infantum* ranges from 29-34% in Brazilian children [58]. Whereas, the prevalence of post-kala-azar dermal leishmaniasis was estimated as 4.4-4.8 per every 10,000 people in India and Bangladesh. The overall number of leishmaniasis cases increased from the last 25 years due to the globalization and the change in climate lead to spread of infection to the non-endemic regions [59].

1.4 VL in India

Since ancient times, VL has been a public health problem in India [60, 61]. It is also called as kala-azar in the Indian subcontinent which mainly affects young children and old people [62]. A total of 7,60,432 VL cases were reported in the duration of 1987-2011. In India, 54 out of 676 districts are affected, that includes 34 districts in the state of Bihar accounting for 70% of the total VL burden on the sub-continent [63]. The total number of annual cases decreased rapidly from 77,102 in 1992 to 12,140 in 2002, with a mean of 30417.28 per year and caused high morbidity and mortality in the years 1992 and 2007 [64]. It is prevalent in 52 districts from four states in India. However, Jharkhand (4 districts), West Bengal (6 districts), Uttar Pradesh (11 districts) and Bihar (31 districts) states were mostly affected [64]. In India *L. donovani* is transmitted anthroponotically by a single species of sandfly vector, *Plebotomus argentipes* [65]. Though asymptomatic individuals are not believed to contribute in transmission, the occurrence of one symptomatic case for every 8.9 asymptomatic cases [66] shows the necessity to elucidate their role in disease transmission [63]. PKDL is appears to be less common in Indian subcontinent, although sufficient data are scarce [67]. HIV infection has not been presumed to be a key factor in the geographic

distribution in India [63]. But, Burza and coauthors reported that 5.6% were positive for HIV among the 2000 adults infected with VL in Bihar [43].

1.5 Transmission

The transmission occurs by the bite of a phlebotomine sandflies infected with *Leishmania* parasites. Upto 98 species belong to the genera *Lutzomyia* and *Phlebotomus* have been proven or suspected as vectors for the human leishmaniasis [68]. Prevalence of sandfly can vary according to the seasons in some geographical areas. The sandfly is 2-3 mm long, noiseless arthropod whose body colour varies from black to white and the wings are positioned at certain angle with respect to abdomen [69]. The sandfly usually takes blood meal in the evening and activity is minimal during the hottest periods of the day [7]. Only the female sand flies which take blood meal from the mammals can support the development of parasite inside the midgut. Some of the sand flies have a wide range of hosts including canids, hyraxes, marsupials and rodents, while others can feed exclusively on humans [70]. According to the transmission patterns, human leishmaniasis can be either zoonotic or anthroponotic [70]. After seeking the blood meal from an infected animal or human, the parasite multiplies within the sandfly gut and then injected into the skin during the subsequent blood meal [35]. Probing the same host for several times is an adaptive method that can increase increases the transmission [71]. Humans with asymptomatic infection, kala-azar or PKDL are the only reservoirs for the anthroponotic transmission [72]. The major reservoir for zoonotic transmission is dog, while in sylvatic cycles that occur in the American rain forests or Central Asia deserts the transmission occurs in enzoonotic mode between wild animals, leaves humans as dead end hosts upon infection [71].

1.6 Life cycle

Leishmania is a digenetic parasite, which completes its lifecycle in two hosts, vertebrate human and invertebrate sandfly of the genus *Phlebotomous* or *Lutzomyia* [72, 73]. The life cycle of the parasite exists between two alternate forms called promastigote and amastigote. Promastigote is an extra cellular, motile, 15-20 μm long and flagellated form which is present in sandfly [74]; and amastigote is intracellular, immotile, 3-5 μm long and non-flagellated that reside inside the mono nuclear phagocytic cells [75]. The life cycle of parasite begins when an infected female sand fly takes a blood meal from healthy individual, it ingests procyclic promastigotes into the blood

stream. *Leishmania* parasite actively invades and multiplies with in the mononuclear phagocytic cells of innate immune system such as dendritic cells, neutrophils and macrophages [76, 77]. The promastigotes actively invade the mammalian macrophage and other mononuclear phagocytic cells and differentiates into amastigotes. Abrupt proliferation of amastigotes results in cell lysis and invasion of surrounding macrophages. In a normal situation, the detection of parasite release recruits inflammatory cells that further stimulate the members of adaptive immune system [10].
When a sand fly probes an infected individual, these amastigotes enters into the sand fly midgut which is enclosed by the type I peritrophic membrane (PM) [78]. The amastigotes differentiates into procyclic promastigotes, proliferates rapidly and differentiates further into nectomonad promastigotes. The nectomonads while moving towards the anterior portion of the gut they differentiate into leptomonad promastigotes and undergoes rapid multiplication and reside the stomodeal valve. After colonization, these leptomonads convert into haptomonads and metacyclics [79]. The metacyclic promastigotes further migrates into the proboscis of sandfly and injects these promastigotes during the subsequent blood meal [80]. The difference in temperature and pH between the insect and mammalian host acts as triggering factors for differentiation of *Leishmania* [81, 82] .

1.7 Treatment

VL was treated from several decades by pentavalent antimonials, which are available as two formulations: sodium meglumine antimoniate and sodium stibogluconate. Increasing non-response to these drugs in India prompted to increased dosage through intravenous infusion or intramuscular injection for over 28-30 days [53]. Due to the resistance to sodium stibogluconate it is longer recommended in Indian subcontinent [83, 84]. Current clinical research has aimed on shorter treatments to avoid resistance [84, 85]. The first advancement has occurred in this direction in India in 2002, with oral miltefosine, followed by paramomysin injection in 2006, for short duration in several combinations, and liposomal amphotericin B (LAMB) for single dose [84]. In east Africa, sodium gluconate in combination with paramomycin is the choice of treatment over a course of 17 days [84, 86]. For all cases of VL caused by *L. donovani* and *L. infantum* in Asia, LAMB surpassed the sodium stibogluconate as the first line treatment. Till date, AmBisome (Gilead Sciences, Dimas, CA, USA) is the only LAMB formulation that was approved by WHO and US Food and Drug Administration (FDA) for the treatment of VL [87]. More than 13000

people have been treated till 2017 with this regimen across the Indian subcontinent under regular program conditions[88].

1.8 Nutrient acquisition of *Leishmania* from hosts

1.8.1 Sand fly

During the lifecycle of *Leishmania*, the parasite is exposed to two different niches which differ in the availability and metabolic adaptation to the nutrients. The diet of sandfly is composed of infrequent blood meals and regular sugar meals which trigger oogenesis [89]. The sand fly acquires the nutrients from the honeydew which contains monosaccharides (fructose, glucose sucrose and maltose), disaccharide (trehalose and trehalulose) and insect synthesized trisaccharide sugars (erlose and melezitose) [90]. *Leishmania* secretes various glycosidases such as chitinase, N-acetyl glucosaminidase, cellulase, α-amylase, α-glucosidase and sucrase which convert the complex sugars to into simple sugars prior to their absorption [91]. Feeding of promastigotes on blood specifies that the parasite should be able to digest amino acids along with the sugars [92]. It was also reported that the promastigotes of *Leishmania* are known to rely on glycolysis and amino acid metabolism for generating energy [92].

1.8.2 Macrophage

After ingestion into the blood stream, promastigotes were internalized by receptor mediated endocytosis and differentiates into the amastigotes in the parasitophorous vacuole (PV) of mammalian macrophage [93]. The PV is an acidic and strongly hydrolytic environment that is rich in cathepsins, glucuronidases, acid phosphatase and Arylsulphatases A and B [94, 95]. Fusion of vesicles derived from the phagosome, autophagosomes, lysosomes and endoplasmic reticulum delivers the complex macromolecules such as glycoproteins, lipoproteins, proteoglycans, which were hydrolyzed to sugars, amino acids, lipids and nucleotides due to the action of hydrolases in the PV [93]. These low molecular weight molecules were transported from the lumen of PV to the cytosol of parasite by specific carriers expressed on the membrane of amastigote and compete with the host transporters [96]. The nutrient transporters are shown to be adapted to acidic milieu in PV with optimum function [96].

1.8.3 Transporters of *Leishmania*

Nutrient import is central event in the lifecycle of parasite for its transmission, to infect host and cause disease [97]. The three glucose transporters, called *Lm*GT1, *Lm*GT2 and *Lm*GT3 have been identified in *L. mexicana* can be expressed both in the promastigote and amastigote stages [98]. One of the isoforms, *Lm*GT2, is upregulated in the promastigote stage may support the higher uptake of glucose [96]. As *Leishmania* are auxotrophic for purines nucleosides, it salvages preformed nucleobases or nucleosides and converts them into phosphorylated nucleotides [99]. The pentose sugar ribose shown to be transported by a specific permease in promastigotes of *Leishmania* [100]. Two purine transporters have been identified in the amastigotes of *L. donovani*, both of which shows high affinity for adenosine [101]. The adenosine/pyramidine transporters are encoded by *Ld*NT1.1 and *Ld*NT1.2, whereas the inosine/guanosine transporter is encoded by *Ld*NT2 [102, 103]. Two additional genes of this family identified in *L. major* designated as NT3 and NT4 were shown to transport adenine, guanine, hypoxanthine and xanthine [104, 105]. Amino acids can be utilized as a carbon source both of the promastigotes and amastigote stages of lifecycle, the uptake of proline is reported in either of the stages in *Leishmania* [106, 107]. In addition, several other important metabolites like polyamines which are required for cell proliferation, myo-inositol for phosphatidyl inositol biosynthesis and biopterin for folate biosynthesis were also shown to be transported by *Leishmania* [108-110].

1.9 Carbohydrate metabolism in *Leishmania*

The carbon metabolism of *Leishmania* is essential for growth and to combat against oxidative stress and microbicidal responses of host [111]. Though knowledge about the nutrient composition of phagolysosome is scarce, increasing evidences revealed that the immune responses of host and activation state of macrophages infected with parasite leads to variability in nutrient composition [112]. *Leishmania* grown in the *in vitro* cultures utilize simple sugars for producing energy (ATP), reducing equivalents and metabolic precursors essential for biosynthesis [111]. As presented in previous section, *Leishmania* imports hexoses, amino sugars and pentoses by expressing various transporters constitutively [113]. Several enzymes involved in the C5 sugar metabolism have been identified in the genome of *L. major,* are ribokinase, ribulokinase and xylulokinase, indicating that *Leishmania* is capable of catabolizing sugars other than hexoses [92]. Studies on *in silico* sequence analysis revealed that the C5-sugar kinases consists of a peroxisome-targeting signal (PTS) which

localizes proteins to glycosomes, besides they facilitate the channeling of C5 sugars into glycolysis *via* the enzymes of pentose phosphate pathway (PPP) [114]. The genes encoding ribulokinase and galactokinase are assumed to be acquired from bacteria by lateral gene transfer since the homologues can be found only in prokaryotes [92].

1.9.1 Glycolysis

Internalized sugars are activated by ATP-dependent sugar-specific (ribose, glucose and galactose) kinases after transporting into glycosomes. For rapid growth, most of the glucose-6-phosphate is channeled into the glycolysis. Analogous to other trypanosomatid parasites, the first seven enzyme of the glycolytic pathway are localized in glycosomes, while the subsequent steps of glycolysis occur in the cytosol [115]. As the early reactions of glycolysis utilize ATP and NAD^+ (for every one mole of Glc-6-P two moles of each), it is necessary to regenerate these molecules in glycosome in order to retain steady-state concentration.

1.9.2 Pentose phosphate pathway (PPP)

The glucose-6-phopshate produced in the glycosomes can also be metabolized through pentose phosphate pathway (PPP) which is connected tightly to glycolysis [116]. The key roles of PPP are to produce reducing power NADPH and other products that enter into the various parts of metabolism. This pathway is particularly important for the parasite survival since the reactive oxygen species produced by the host as a first-line antimicrobial defense counteracted by NADPH dependent reactions [116]. In *Leishmania* and other trypansomatid parasites this is achieved by import and subsequent conversion of phospho enol pyruvate (PEP) to succinate or decarboxylation to pyruvate [111]. The fermentation of succinate in glycosomes results in the production of C4 dicarboxylic acids which enter into TCA cycle and participate in anaplerotic reactions [117]. While, the glycerol-3-phosphate/dihydroxyacetone phosphate shuttle occur between the glycosomes and mitochondria balances the NAD^+ / NADH levels [118].

1.9.3 Mannogen synthesis

Other than glycolysis and PPP, the excess hexose-6-phosphates may be transported to cytosol for incorporation into mannogen, a short term carbohydrate reserve [119]. Mannogen comprises of short chains (4-40 residues) of mannose residues connected by β-1,4 linkage that accumulates in

intracellular amastigotes and stationary phase promastigotes [119]. On the other hand, *Leishmania* is capable of synthesizing mannogen from a glucogenic precursor, L-lactate [120].

1.9.4 Gluconeogenesis

The capacity to uptake glucose is considerably less in amastigotes compared to promastigotes [121]. Throughout their life cycle, *Leishmania* parasites import and assimilate glucose and other sugars [121, 122]. The promastigotes take up sugar from the extracellular milieu upon its availability, and also able to synthesize sugar under starved condition *via* gluconeogenesis. In contrast, amastigotes actively execute gluconeogenesis irrespective of sugar in the environment, suggesting the adaptation of amastigote to intracellular environment that is sugar-limited [120]. It has been reported in *L. mexicana* that the enzymes pyruvate phosphate dikinase (PPDK), glycerol kinase (GK) and phosphoenol pyruvate carboxy kinase (PEPCK) enable the entry of alanine, glycerol and aspartate into the gluconeogenesis, respectively [120].

1.10 Ribose metabolism in *Leishmania*

D-Ribose is an abundant pentose sugar in all living organisms including prokaryotes and eukaryotes. The role of ribose inside cell is critical with respect to its structure and metabolism (Fig. 1). As part of the ATP and GTP it plays a critical role in the cellular energy currency [123]. As a constituent of sugar-phosphate backbone of RNA and DNA, it plays both structural and informational roles [123]. In addition, ribose is one of components of several nucleotide derived co-factors. The honey dew on which the sand fly vector of *Leishmania* feed upon known to contain ribose along with other sugars [124]. Free ribose likely to be available in large quantities to amastigotes as a consequence of cytoplasmic RNA turnover through autophagosomal pathway

followed by catabolism in endosomes [125]. *Leishmania* imports ribose from the environment by expressing specific transporters.

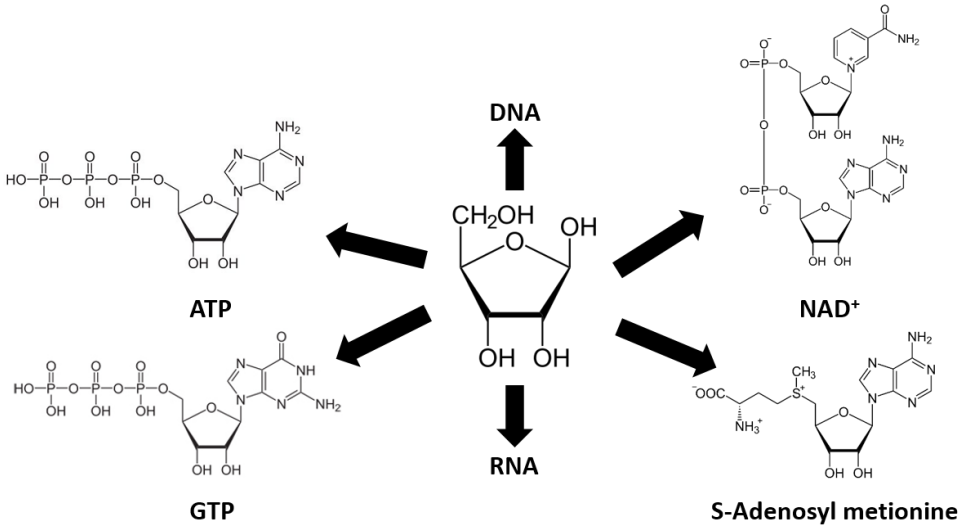


Figure 1. Significance of ribose

1.10.1 Transport of ribose by specific membrane transporter

Ribose plays an essential role to support the promastigote growth under glucose lacking conditions [126]. As indicated previously, free ribose is likely to be available both in alimentary tract of sandfly and parasitophorous vacuole of mammalian macrophage. A permease mediated ribose uptake was reported from promastigotes of *L. donovani* with an apparent *K*m of 2 m*M* [127]. The transport of ribose is independent from the hexose transport system that was well characterized in *Leishmania* since ribose was not able to obstruct the uptake of glucose or fructose [100, 128]. After transport, the ribose pool in the promastigotes consists of one third in un-phosphorylated form and two thirds in phosphorylated form [100]. Further, the neutral form of the ribose was exchanged with the exogenously added 2'-deoxy ribose. Similarly, a carrier mediated ribose accumulation has also been reported in *L. mexicana* [128]. Facilitating ribose entry through a transporter may provide *Leishmania* with ribose moiety required for purine and pyrimidine nucleotide biosynthesis.

1.10.2 Uptake of ribose by GLUT

In *L. mexicana*, three glucose transporters (*Lm*GT1, *Lm*GT2 and *Lm*GT3) have been identified and characterized till date [129]. Among them, *Lm*GT1 was identified as a most divergent member with low relative affinity for glucose. While the other two members of *Lm*GT family share identity greater than 90%, and with similar affinity for glucose [130]. Though both of these transporters are responsible for glucose uptake, *Lm*GT2 was up-regulated in the promastigote stage and showed higher affinity and capacity to transport ribose [130]. The relative specificity (V_{max} / K_m) of *Lm*GT2 for ribose is six folds higher than that of *Lm*GT3. A pair of alanine residues located in the loops joining helices 3-4 and 7-8 in *Lm*GT2 acts as a selectivity filter [130]. The substitution of alanine pair with threonine residue pair in *Lm*GT3 limited the binding of ribose to exofacial substrate binding pocket. It was also reported that the Δ*lmgt* promastigotes are not able to uptake D-ribose in addition to glucose [113, 130]. A fourth transport protein, *Lm*GT4, which is a divergent permease transports hexose sugars with low affinity and overexpressed upon deletion of *Lm*GT1-*Lm*GT3 [131]. These findings revealed the most robust transport of ribose by a specific glucose transporter isoform (*L*mGT2) in *Leishmania*.

1.10.3 Hydrolysis of nucleosides by nucleoside hydrolases

Leishmania parasites are auxotrophic for purines due to the lack of *de novo* pathways for purine biosynthesis [132, 133]. They depend on host derived purine bases or nucleosides in order to meet growth demands. Among the enzymes implicated in purine salvage, a cell surface bound 3'-nucleotidase / nuclease is anticipated to produce nucleosides [134]. Nucleoside N-hydrolases (NHs) play crucial role in the salvage of purines by hydrolyzing N-glycosidic bond of pyramidine and purine nucleosides to produce free ribose and bases. From *L. donovani*, three different nucleoside hydrolases have been purified [134]. Among these, two are ribonucleosidases that exhibit different preferences for substrate: one is purine and pyramidine nucleosidase; the other one cleaves only purines. One of the NH from *L. donov*ani has been identified as non-specific nucleoside hydrolase based on significant activity on natural purine and pyramidine nucleosides, and p-nitrophenyl-β-D-ribofuranoside [127]. Several nucleoside hydrolases having different substrate specificity present in a single species of *Leishmania* generates high intracellular ribose pool.

1.11 Production of ribose-5-phosphate in *Leishmania*

Ribose is utilized in various metabolic pathways in the form of ribose-5-phosphate (R5P). The metabolic pathways involved in the generation of ribose-5-phosphate have a great importance due to its role as a precursor for many biological molecules including nucleotides, nucleic acids (RNA and DNA), nucleotide based co-factors (FAD and coenzyme A) and amino acid histidine [128, 135, 136]. A nucleotide is composed of nitrogen base, ribose sugar and a phosphate moiety. The ribose moiety is incorporated into nucleotide after activation into the 5-phosphoribosyl-1-pyrophosphate (PRPP) in the presence of PRPP synthetase (PRS) [137]. Ribose-5-phophate is produced in *Leishmania* by one or combination of pathways mentioned as below. These pathways may also converge in order to produce ribose-5-phopshate for nucleic acid biosynthesis [128].

1.11.1 Production of ribose-5-phosphate through pentose phosphate pathway

The pentose phosphate pathway comprises of two phases, oxidative phase and non-oxidative phase (Fig. 2). The oxidative phase comprises of three reactions which convert one mole of glucose-6-phosphate (G6P) into one mole of ribulose-5-phosphate (Ru5P), one mole of CO_2 and also generates two moles of NADPH [116]. Usually the flux of G6P through PPP depends on the ratio of NADPH : NADP [138], requirement for ATP and R5P. Ru5P is the final product of the oxidative phase, and enters into the non-oxidative phase [116]. Subsequently, the Ru5P is converted to either R5P by ribose-5-phophate isomerase, a critical metabolite used for the nucleotide synthesis, or to xylulose 5-phosphate (X5P) in the presence of ribulose-5-phosphate epimerase [116].

Both of the products, R5P and X5P acts as substrates for transketolase (TKT), that transfers two carbon unit to an aldose from a ketose to produce glyceraldehyde 3-phosphate (GA3P) and sedoheptulose 7-phosphate (S7P) [116]. These are used further by transaldolase that transfers three carbon unit to GA3P from S7P to produce fructose 6-phosphate (F6P) and erythrose 4-phosphate (E4P). The former re-enters glycolysis and E4P is used in TKT reaction along with X5P that results in production of GA3P and F6P. Transketolase (TKT) can also catalyze the reverse reaction by condensing F6P and GA3P to regenerate R5P and X5P [116]. A metabolomic studies from *tkt* null cell line of *L. mexicana* (*lmtkt* - / -) revealed an increase in R5P level [139]. The phosphorylated sugars produced during the various stages of the pathway can be channeled into

glycolysis or used as a precursors for the biosynthetic reactions, erythrose-4-phosphate is for the aromatic amino acid biosynthesis.

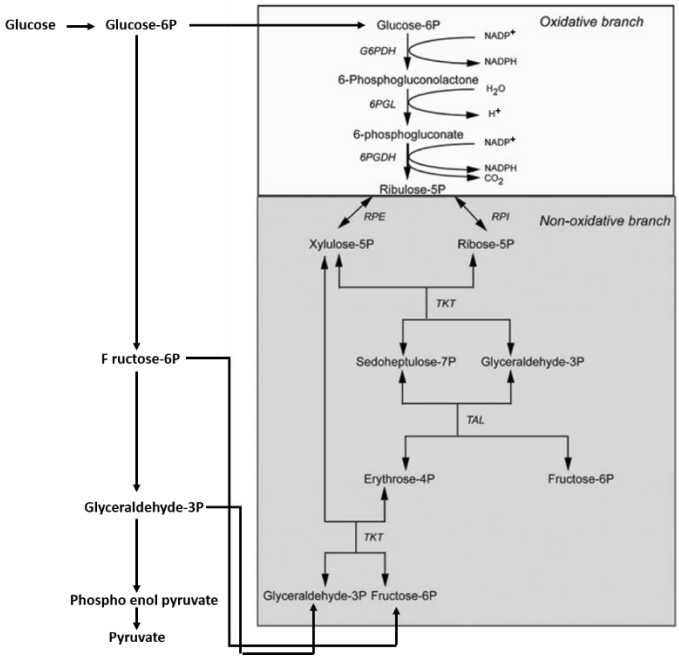


Figure 2. Pentose phosphate pathway. Metabolism of glucose through glycolysis and through pentose phosphate pathway by branching at glucose 6-phosphate. **(*Adapted and modified from Anais da Academia Brasileira de Ciências, 79(4), 649-663, 2007*).**

1.11.2 Generation of ribose-5-phosphate from ribose

The PPP is not the only potential source for R5P in *Leishmania* as this metabolite can also be produced from ribose in the presence of ribokinase [128]. As discussed in the section 6.1, 6.2 and 6.3, free ribose has been shown to be accumulated in the parasite cytosol *via* a specific carrier, GLUT and also by cleavage of nucleosides by hydrolases. These results had shown the existence of multiple pathways in generation of the single metabolite, ribose-5-phosphate in *Leishmania* (Fig. 3). Phosphorylation plays a crucial role in the metabolism of monosaccharides [140]. Before entering into the metabolic pathways, ribose should be activated by phosphorylation, which is the first step in the metabolism of D-ribose. Un-phosphorylated ribose cannot be channeled into the major metabolic pathways in any of the organisms which can utilize ribose in its metabolism [136].

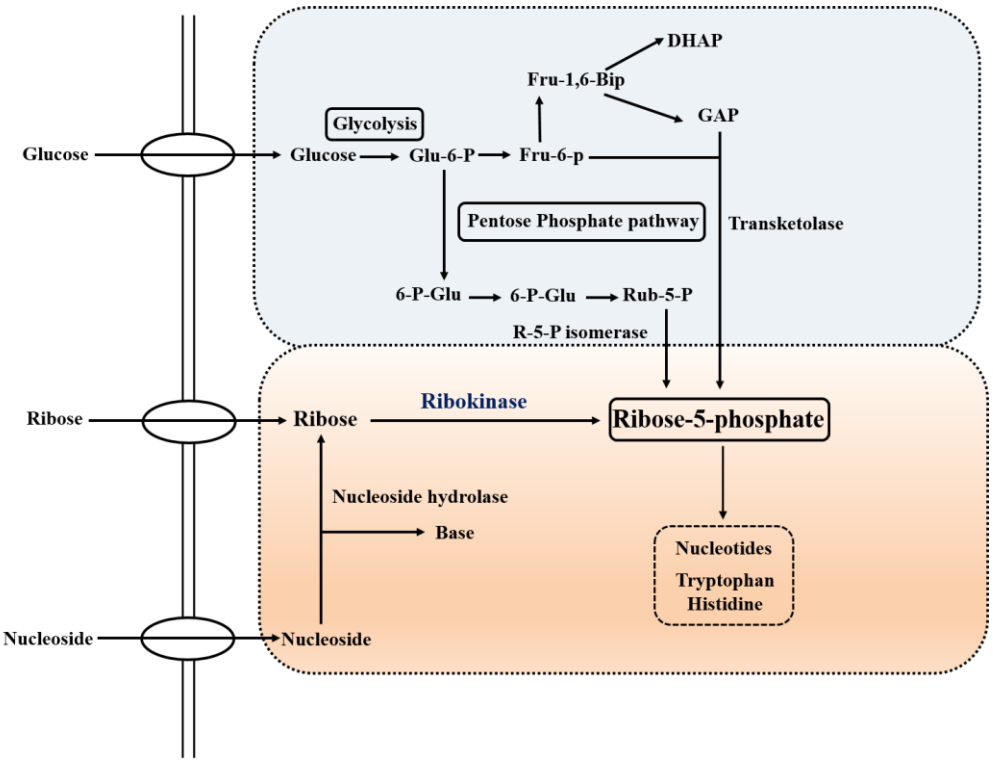


Figure 3. Pathways involved in the production of ribose-5-phosphate in *Leishmania*. The pathways that contribute in the generation of ribose 5-phosphate pool to synthesize nucleic acids are shown here. Modified from **(*Adapted and modified from Molecular and biochemical parasitology, 130(2), 117-125, 2003*).**

1.11.3 Role of kinases in metabolism

Kinases are one of the most extensively studied enzymes which catalyze the phosphate group transfer from a donor to an acceptor molecule. Compared to other members of the group, protein kinases are the most extensively explored members compared to those involves in the metabolism. RK-like kinases phosphorylates a variety of metabolites such as carbohydrates, nucleotides, vitamins and other cofactors mostly at hydroxyl methyl group [141]. Even though the sequence identity among these members is quite low, they share similar three dimensional architecture with a nucleotide binding motif and broad substrate specificity [142].

Comparative sequence analysis of ribokinase genes from both prokaryotes and eukaryotes has shown that they belong to the carbohydrate kinases of PfkB family, also called as RK family [143]. The members of ribokinase family catalyze the transfer of γ-phosphate from ATP to the hydroxyl

methyl groups of a variety of the sugars. Ribokinase family along with two other families of proteins called, hexokinase and galacto kinase were originated by convergent evolution [143].

1.11.4 Role of RK in metabolism of *Leishmania*

The free ribose accumulated by various metabolic pathways gets activated by ribokinase and subsequently enters into the pentose phosphate pathway for energy production or gets utilized for the biosynthesis of nucleotides and amino acids such as tryptophan and histidine [144, 145]. Addition of a negatively charged phosphate group reduces the permeability of ribose across the lipid bilayer [140]. Experiments with radiolabeled glucose and ribose in *L. mexicana* have shown the relative incorporation of glucose and free ribose into nucleic acids after activation by RK. It was shown that about 10.9 % of D-glucose is used for nucleic acid synthesis [128], which shows the flux of D-glucose through oxidative phase of PPP is 10%. Whereas, 56.7% of D-ribose used for nucleic acid incorporation suggests that 43.3 % of D- ribose may have fluxed through the non-oxidative branch of PPP to produce glycolytic intermediates [128].

Under normal conditions in *Leishmania*, the majority of ribose-5-phosphate gets produced from glucose *via* pentose phosphate pathway [128], whereas, the glucose-limited state in *Δlmgt* promastigotes revealed the production of ribose-5-phosphate by phosphorylation of free ribose by up-regulation of non-specific nucleoside hydrolase and RK [146]. Nevertheless, the *Δlmgt* promastigotes seem to degrade nucleotides to salvage ribose and recycle for new nucleotide synthesis [146]. These studies show the significance of ribokinase in the metabolism of *Leishmania* when the availability of hexose sugars is limited or absent. Also, ribokinase is predicted to play a crucial role in maintaining the ATP-ADP balance in glycosomes in the related trypanosomatid parasite *Trypanosoma brucei* by catalyzing the reverse reaction [147]. *In silico* sequence analysis shown that RK from *L. major* predicted to contain a peroxysomal targeting sequences 2 (PTS2) at the N-terminus, which supports that RK feeds substrates (R5P) directly into PPP in glycosomes [114].

It is well known that throughout the life cycle *Leishmania* encounters ribose. Among the three major pathways that produce ribose-5-phosphate in *L. donovani*, RK plays a critical role by phosphorylating the free ribose to enable its entry into metabolic pathways [128]. The structure based functional analysis of ribokinase was extensively investigated from the prokaryotes such as

Escherichia coli, *Staphylococcus aureus and Vibrio cholera* [148-151]. Whereas, in case of eukaryotes specifically the knowledge about RK from trypanosomatids is scarce. Even though the detailed mechanism of ribose production in the metabolism of *Leishmania* was explored, the structural understanding of RK catalysis is not yet elucidated. Based on these preliminary studies on the role of ribokinase in the metabolism of *Leishmania*, following objectives were proposed to gain insight into the catalytic mechanism of *Ld*RK:

1.12 Objectives

Objective 1

Molecular cloning and purification of ribokinase from Leishmania donovani (LdRK).

Objective 2

Biophysical and biochemical characterization of recombinant LdRK.

Objective 3

Crystallization and comparative structural analysis of apo and ligand bound structures of LdRK.

2. Materials and Methods

2.1 Materials

2.1.1 Plasmids and bacterial strains used for cloning and expression

The cloning vector pJET1.2 was purchased from MBI Fermentas and the expression vector pET-28a was obtained from Novagen, Madison, USA. Cloning and expression was performed using the bacterial strains *E. coli* XL1-Blue (Stratagene, USA) and *E. coli* BL21 (DE3) (Novagen), respectively.

2.1.2 Culture media and antibiotic stocks

2.1.2.1 Luria-Bertani (LB) medium

For preparing LB medium, 10 g of tryptone, 10 g of NaCl and 10 g of yeast extract were dissolved in 900 ml of water and volume was adjusted to 1000 ml.

2.1.2.2 Antibiotic stocks

Ampicillin (100 mg/ml) and kanamycin (50 mg/ml) stocks were prepared in sterile milliQ water, while tetracyclin (12.4 mg/ml) was prepared in 70 % ethanol and stored at -20˚C. The components of culture media and antibiotics were purchased from HiMedia Laboratories (India).

2.1.3 Molecular biology

The enzymes *Pfu* DNA polymerase, *Nde*I and *Xho*I and kits including GeneJET Gel Extraction Kit and GeneJET Plasmid Miniprep Kit were purchased from Thermo scientific (USA). Primers used for the amplification were purchased from IDT Integrated DNA Technologies, Inc.

2.1.4 Enzymes assays

All enzyme assay components such as pyruvate kinase (PK), lactate dehydrogenase (LDH), ribose, ATP, phospho enol pyruvate, NADH were obtained from Sigma Aldrich (München, Germany). The monovalent salts (NaCl, KCl, CsCl, K_2HPO_4), $KHCO_3$ and $HClO_4$ were purchased from HiMedia Laboratories (India) and divalent salts ($MgCl_2$, $MnCl_2$, $CoCl_2$, $CaCl_2$, and $NiCl_2$) were purchased from Merck specialties, India.

2.1.5 Crystallization

Crystallization accessories were purchased from Molecular Dimensions (UK). The components of reservoir solution, citric acid, NaCl and glycerol as well as the ligands (ATP, AMPPCP, ADP and ribose) used for co-crystallization were procured from Sigma Aldrich. Silicone oil was purchased from Merck specialties, India.

2.2 Methodology

2.2.1 Cloning

The full-length ORF for ribokinase was amplified from the genomic DNA of *L. donovani* by polymerase chain reaction using the oligonucleotide primers (Table 1). The amplified PCR product was purified using a gel-extraction kit and ligated into pJET1.2 cloning vector. The insert was released from the pJET1.2 vector by digestion with the *Nde* I and *Xho* I restriction enzymes and was then subcloned into pET-28a(+) expression vector, which adds a hexa-histidine tag to the N-terminus of the target protein. The plasmids isolated from the positive clones were verified by restriction digestion followed by sequencing prior to transformation into *E. coli* BL21 (DE3) cells.

Table 1: Primers used for the amplification

*Ld*RBK-F	5' AATCAT ATGCACCGTGCGCGGAACGTT3'
*Ld*RBK-R	5' AATCTCGAGCTACATGACACCAGCCGGCA3'

2.2.2 Expression and purification

For overexpression, a single colony was inoculated into 30 ml Luria-Bertani (LB) medium containing 50 mg $m1^{-1}$ kanamycin and incubated overnight at 310 K. An overnight-grown culture (20 ml) was diluted to 1% with LB broth, incubated at 310 K to an OD_{600} of 0.6 and then induced with 0.05 m*M* IPTG for 20 hrs at 291 K. Induced cells were pelleted for 10 min at 6,300g and resuspended in 50 ml lysis buffer (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 40 m*M* imidazole, 2 m*M* β-mercaptoethanol, 0.25 mg ml^{-1} lysozyme, 0.1 mM phenyl methyl sulfonyl fluoride) followed by lysis using sonication. The cell lysate was centrifuged at 21,000g for 30 min, and the supernatant was filtered using a 0.2 mm filter (Millipore) and loaded onto a HisTrap HP 5 ml column (GE Healthcare) equilibrated with buffer A consisting of 50 m*M* Tris–HCl pH 8.0, 300

m*M* NaCl. Unbound bacterial proteins were eluted with a step gradient of 10, 15, 20 and 30% buffer B (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 500 m*M* imidazole) and the desired protein was eluted from the column using 80% buffer B. Eluted fractions with high purity were pooled and loaded onto a HiLoad 16/60 Superdex 200 pg column (GE Healthcare) equilibrated with 50 m*M* Tris–HCl pH 8.0 and 125 m*M* NaCl. The fractions with the highest purity were pooled and then concentrated using Amicon Ultra (10 kDa cutoff) to 8 mg ml^{-1} for crystallization. All purification steps were performed at 277 K and protein concentrations were determined at 280 nm with a NanoDrop 2000 (Thermo Scientific) using a molar extinction coefficient (ε) of 16390 M^{-1} cm^{-1} and the molecular weight of ~37.5 kDa.

2.2.3 Circular dichroism

The circular-dichroism (CD) spectrum was measured using 7.5 m*M* Tris–HCl pH 8.5 containing 3.2 m*M* *Ld*RK with a J-1500 spectropolarimeter (Jasco) at 298 K from 260 to 190 nm with a scanning speed of nm min^{-1}. The path length of the sample cell was 2 mm, while the data interval and bandwidth were both 1 nm. Two consecutive spectra were recorded and averaged, and the data were plotted after subtracting the buffer blank in units of mdeg. Secondary-structure content was evaluated using the *CDSSTR* algorithm with reference data set No. 4 from the *DichroWeb* server (Whitmore & Wallace, 2008). Thermal stability was monitored in 10 m*M* sodium phosphate buffer pH 7.4 by measuring the ellipticity at 222 nm from 293 to 343 K with a temperature gradient 1 K min^{-1}. The melting temperature (*T*m) was calculated by plotting the first derivative of the thermal denaturation curve. All spectra were smoothed before plotting.

2.2.4 Enzyme assay

Enzyme assays were performed in triplicate at 298 K in 96-well plate using microplate reader (Molecular Devices) with path length 0.596 cm. Activity of ribokinase was measured with the method reported previously [152] with some modifications by coupling the production of ADP to the oxidation of NADH and measuring the change in absorbance at 340 nm. The reaction mixture contains 30 m*M* Tris-HCl pH 8.5, 100 m*M* KCl, 3 m*M* MgCl2, 0.3 m*M* PEP, 0.2 m*M* NADH, pyruvate kinase (0.4U) and lactate dehydrogenase (1.2U) as coupling enzymes. For determination of initial reaction rates, the concentration of ATP was fixed at 0.8 m*M* and ribose was varied from 0.01 to 4 m*M*, whereas the concentration of ribose was fixed at 4 m*M* and ATP was varied from

0.005 to 0.8 m*M*. The reactions were started by addition of 0.2 µ*M* enzyme and absorbance was recorded at every 10 sec for 30 min. The kinetic data was fitted to the Michaelis-Menten equation through non-linear regression using *GraphPad Prism*.

2.2.5 Effect of monovalent and divalent cations

The monovalent cation activation assay was performed separately for each monovalent cation salt (CsCl, KCl and NaCl) at a fixed 100 m*M* concentration in a standard reaction mixture containing 30 m*M* Tris–HCl (pH 8.5), 3 m*M* $MgCl_2$, 4 m*M* ribose and 0.8 m*M* ATP. Whereas for divalent cation effect assay, 3 m*M* concentration of each divalent cation salt ($MgCl_2$, $MnCl_2$, $CoCl_2$, $CaCl_2$, $CuCl_2$, and $NiCl_2$) was added separately in a standard reaction mixture containing 30 m*M* Tris–HCl (pH 8.5), 100 m*M* KCl, 4 m*M* ribose and 0.8 m*M* ATP.

The assay was initiated by adding 0.2 µ*M* of *Ld*RK to a 0.5 ml reaction system followed by incubation at 25°C. After 5 min, the enzyme activity was stopped by adding $HClO_4$ to the final concentration of 3% and then incubated on ice for 5 min. Subsequent to neutralization with $KHCO_3$, the tube was centrifuged at 10,000g for 20 min at 4°C and supernatant was added to the coupled assay reaction mixture containing 100 m*M* KCl, 2 m*M* $MgCl_2$, 0.3 m*M* phosphoenolpyruvate, 0.2 m*M* NADH, 0.4 U pyruvate kinase and 1.2 U lactate dehydrogenase. The reaction volume was adjusted to 0.7 ml and the amount of ADP that was produced by the activity of *Ld*RK was assessed by measuring the oxidation of NADH at 340 nm.

2.2.6 Crystallization

2.2.6.1 Screening and crystallization of *Ld*RK-apo

Crystallization screening was carried out initially by sitting drop vapor diffusion method at 296 K using a Mosquito Protein Crystallization Robot (TTP Labtech at IICT, Hyderabad) in MRC 96-well plates by mixing 0.15 µl of protein (*Ld*RK) with 0.15 µl of precipitant solution. Commercially available crystallization kits from Hampton Research (Index, PEG ion, SaltRx, Crystal screen, Crystal screen 2), Molecular Dimensions (Structure screen 1, Structure screen 2, Morpheus, Midas, PGALM, PACT premier, JCSG plus) were used for the screening of suitable condition. Hexagonal rod-shaped crystals appeared after one day in a condition containing 0.1 *M* citric acid pH 3.5 and 3.0 *M* NaCl. Further optimization was performed using hanging-drop vapour-diffusion method by

varying concentration of protein, pH of the buffer and ionic strength of NaCl in 24-well plates with a reservoir volume of 500 µl. Crystals suitable for the diffraction experiments were obtained after 21 days with 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl and 4.5% (v/v) glycerol (Table 2). Simultaneously to confirm the presence of RK, crystals were washed with well solution, dissolved in sample loading buffer, electrophoresed on SDS-PAGE followed by silver staining.

2.2.6.2 Co-crystallization of *Ld*RK with ATP, AMPPCP and ADP

The diffraction quality crystals obtained with an optimized condition composed of 0.1 *M* citric acid (pH 4.3), 3.4 *M* NaCl and 4.5% glycerol was used further to optimize the co-crystallization with nucleotides. To eliminate the interference due to the acidic nature of ATP during crystallization, the ATP stock solution was prepared using 0.1 *M* Tris-HCl (pH 9.0) that brings the pH of ATP solution nearer to neutral. The ATP complex crystals were grown by mixing reservoir solution containing 0.1 *M* citric acid (pH 4.3), 3.4 *M* NaCl, 10 m*M* $MgCl_2$ and 10% glycerol with an equal volume of *Ld*RK (3.8 mg/ml) incubated with 2.5 m*M* ribose and 5 m*M* ATP. Crystals of AMPPCP complex were grown using the same condition by incubating the protein with 5 m*M* AMPPCP instead of ATP before crystallization. In co-crystallization attempts with the products such as ribose-5-phosphate and ADP, the size of the crystals got reduced. To improve the size of the crystals, streak seeding was carried out along with controlled vapour diffusion using silicone oil. The crystals obtained initially with the reservoir solution containing 0.1 *M* citric acid (pH 4.3), 3.4 *M* NaCl, 10 m*M* $MgCl_2$ and 9% glycerol were used for the preparation of seed stock. The seed stock was then diluted to 1:200 with the reservoir solution, streaked over the drop and equilibrated against the reservoir solution layered with 250 µl of silicone oil.

2.2.6.3 Data processing and structure determination

Diffraction data were collected at the Indus-2 beamline BL21 of the Raja Ramanna Centre for Advanced Technology (RRCAT, Indore, India) using a MAR 225 CCD detector for the apo, ATP and ADP complex crystals, and a Rigaku Micromax 007 with R-axis IV++ detector system at Indian Institute of Chemical Technology (IICT, Hyderabad, India) for AMPPCP complex crystals. Data processing and reduction were carried out with *XDS* [153]/ *HKL2000* [154]. The data collection and refinement statistics are summarized in Table 2.

Chain A of human RK (PDB: 2FV7) was used as a search model for molecular replacement with *MOLREP* [155] to find out the initial phases for *Ld*RK-ATP complex. After that, the chain B of *Ld*RK-ATP complex was used to determine the structures of apo, ADP and AMPPCP complex. The structures were subjected to iterative cycles of model building with *COOT* [156] followed by refinement using *REFMAC5* [157] from *CCP4* suite. The quality of the final structures was checked by *PROCHECK* [158]. The atomic coordinates for the apo, ATP, AMPPCP and ADP complex structures have been deposited to PDB with the accession IDs 5ZWY, 6A8A, 6A8B and 6A8C, respectively.

Table 2: Data collection and refinement statistics for structures of *Ld*RK

Parameter	***Ld*RK-apo**	***Ld*RK-ATP**	***Ld*RK-AMPPCP**	***Ld*RK-ADP**
Data Collection				
Beamline	RRCAT-BL21	RRCAT-BL21	Rigaku MicroMax-007 HF rotating anode	RRCAT-BL21
Detector type	Marmosaic 225 mm CCD	Marmosaic 225 mm CCD	R-axis IV++ detector system	Marmosaic 225 mm CCD
Wavelength (Å)	0.9794	0.9794	1.54	0.9794
Data collection temperature (K)	100	100	100	100
Space group	$P6_1$	$P6_1$	$P6_1$	$P6_1$
a, b, c (Å)	100.25, 100.25, 126.77	99.90, 99.90, 125.55	100.42, 100.42, 125.79	100.70, 100.70, 126.28
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution (Å)	46.61-1.95 (2.00-1.95)	46.41-1.80 (1.84-1.80)	32.87-2.01 (2.06-2.01)	51.14-1.98 (2.03-1.98)
R_{merge} (all *I*+ and *I*-)	0.101 (0.604)	0.070 (0.824)	0.121 (0.554)	0.139 (0.537)
Mean $I/\sigma I$	14.9 (3.4)	17.9 (2.1)	8.2 (3.1)	10.7 (3.9)
CC(1/2) (%)	99.8 (87.6)	99.6 (82.0)	59.9 (63.2)	99.5 (89.8)
Total number of reflections	395500	519604	244361	388772
Mosaicity (°)	0.09	0.19	0.60	0.30
Completeness (%)	100 (99.9)	100 (100)	99.8 (100)	99.9 (99.9)
Multiplicity	7.5 (7.5)	7.9 (6.0)	5.1 (5.0)	7.7 (7.7)
Wilson B-factor ($Å^2$)	16.7	21.9	27.4	12.0
Matthews coefficient ($Å^3$/Da)	2.45	2.41	2.44	2.46
Solvent content (%)	49.87	49.03	49.64	50.12
No. of molecules in ASU	2	2	2	2
Refinement				
No. of unique reflections	52575	65706	47585	50454
R_{work}/R_{free} (%)	19.2/22.5	18.9/21.2	20.0/24.0	17.9/21.6
Total no. of non-H atoms	5030	5286	5158	5375
No. of water molecules	204	243	158	296
Mean B-factor ($Å^2$)	28.2	31.4	34.7	21.7
R.m.s. deviations: Bond lengths(Å) Bond angles (°)	 0.01 1.52	 0.01 1.73	 0.01 1.15	 0.01 1.27
Ramachandran favoured (%)	98	97	97	98
Ramachandran allowed (%)	2	3	3	2
Ramachandran outliers (%)	0	0	0	0
PDB id	5ZWY	6A8A	6A8B	6A8C

Chapter I

3. Molecular cloning and purification of ribokinase from *Leishmania donovani* (*Ld*RK)

3.1 Results

3.1.1 Multiple sequence alignment

To identify the motifs involved in the binding of substrates (ribose and ATP) and other signature sequences, *L. donovani* RK was aligned with RKs sequences from *E. coli, S. aureus, V. cholerae* and *H. sapiens,* whose structures are already known. The sequence of *Leishmania* RK was more similar to human protein (40% identity) than to the aligned bacterial RK sequences. Insertion of a five amino acid stretch at two of the positions i.e. 143-148 (-NYERI-) and 190-194 (- KPAEV-) was observed exclusively in the amino acid sequence of *Ld*RK (Fig. 4). The primary structure of *Ld*RK contains the - NEVE - (209-212) and -GAGD- (275-278), which are the two signature sequences of proteins belonging to the PfkB family [159-161]. The - GAGD - motif contains the catalytic base, aspartic acid (D278) whereas the other three residues (G275, A276 and G277) binds to γ-phosphate of ATP and stabilizes the transition state during catalysis [159]. Most of the residues involved in the binding of ribose and ATP were conserved with other RKs. Out of seven residues that form direct hydrogen bonds with ribose, only F25 of *Ld*RK was replaced with asparagine in *Ec*RK, *Sa*RK, *Vc*RK, and methionine in *Hs*RK [150, 159] . Furthermore, N302 of *Ld*RK was aligned with histidine residue that makes interaction with ribose moiety of ATP in *Ec*RK and *Vc*RK [150, 159]. Residues involved in monovalent cation binding were also found to be conserved with other RKs.

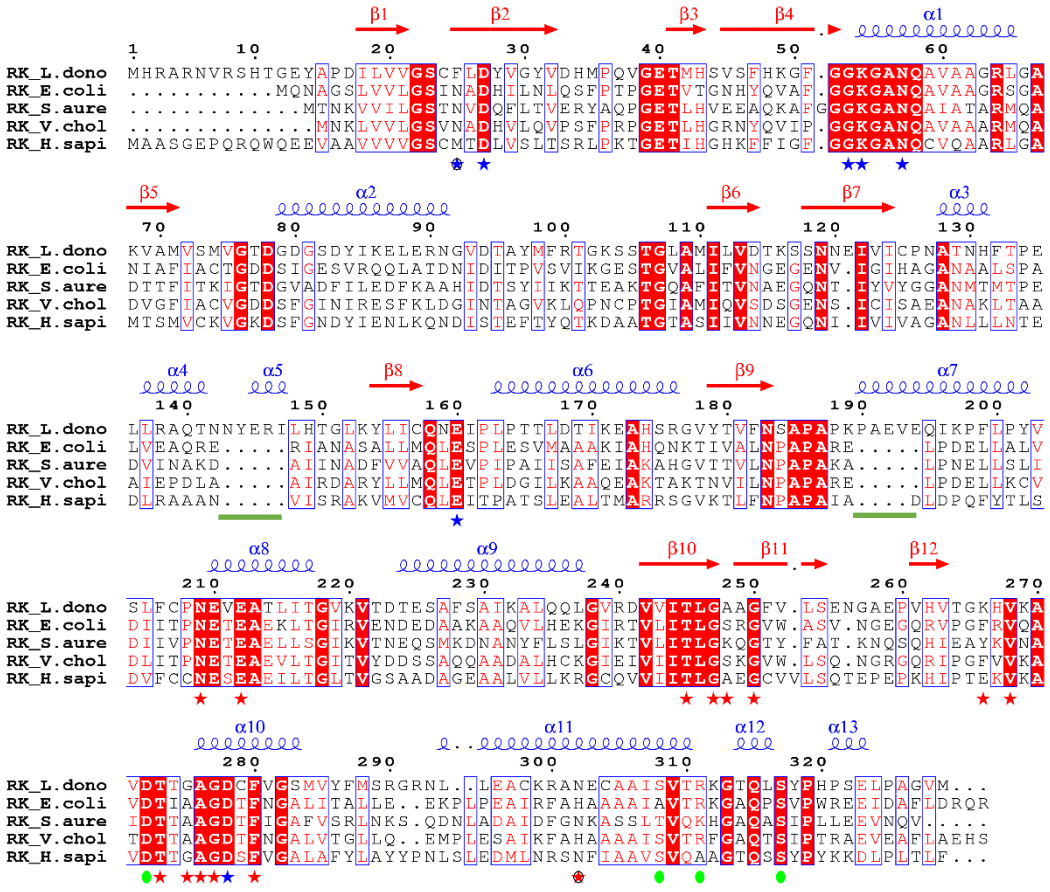
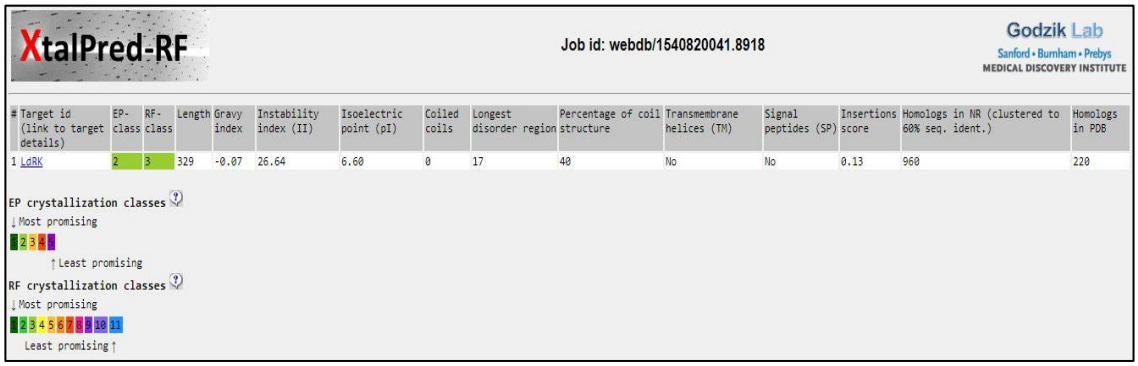


Figure 4. Multiple sequence alignment of *Ld*RK with its counterpart from *Escherichia coli, Vibrio cholerae, Staphylococcus aureus* and *Homo sapiens*. The secondary structural elements, α-helices and β-sheets are shown in blue and red, respectively. Green lines specify the insertion of a five amino acid stretch observed at the positions, 143-147 and 190-194. Monovalent cation binding residues are depicted by green circles. The proposed ribose binding residues are indicated with blue stars whereas the red stars indicate residues involved in the binding of ATP. Circled stars designate F25 and N302 that are observed in the ribose and ATP pockets of *Ld*RK, respectively. Sequence alignment was performed with Clustal Omega and figure was generated with *ESPript* (http://espript.ibcp.fr/ESPript/ESPript/).

3.1.2 Crystallizability prediction of *Ld*RK

Before proceeding for cloning, the crystallizability of *Ld*RK was predicted using *XtalPred* web server [162]. The full-length (329) construct was classified to have "sub optimal" chance of crystallizability with an expert pool and random forest crystallizability classes 2 and 3, respectively (Fig. 5). No disordered regions, transmembrane helices or signal sequences are predicted in the protein sequence of *Ld*RK according to outcome of *XtalPred.*



#	Target id (link to target details)	EP-class	RF-class	Length	Gravy index	Instability index (II)	Isoelectric point (pI)	Coiled coils	Longest disorder region	Percentage of coil structure	Transmembrane helices (TM)	Signal peptides (SP)	Insertions score	Homologs in NR (clustered to 60% seq. ident.)	Homologs in PDB
1	LdRK	2	3	329	-0.07	26.64	6.60	0	17	40	No	No	0.13	960	220

Figure 5. Prediction of *Ld*RK crystallizability. The sequence based crystallizability prediction of *Ld*RK is most promising according to *XtalPred* with the EP-class score 3 and RF-class score 4.

3.1.3 Cloning and expression of *Ld*RK

The ORF encoding the full-length RK was PCR-amplified from genomic DNA of *L. donovani* DD8 using gene specific primers (Fig. 6A). The amplicon was cloned initially into pJET1.2 cloning vector. After identifying the positive clones by colony PCR the plasmid isolated from one of positive clones was confirmed through restriction digestion (Fig. 6B). Followed by, the insert released from pJET vector was sub cloned into pET-28a (+) expression vector. Plasmid isolated from a positive clone was verified by restriction digestion and the orientation and sequence of ORF was confirmed through DNA sequencing (Fig. 6C). The *Ld*RK-pET28a construct was then transformed into *E. coli* BL21 (DE3) expression cells. Subsequent solubility analysis has shown that approximately 70 % of the recombinant protein was obtained in the soluble fraction.

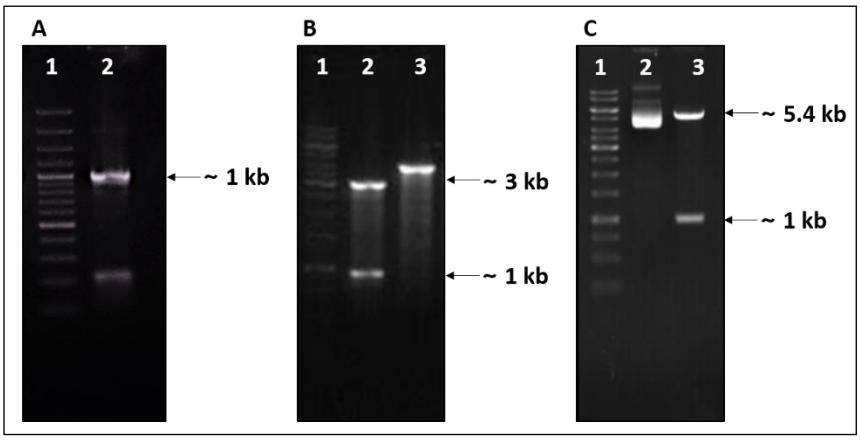


Figure 6. Cloning of *Ld*RK into pET-28a expression vector. (A) PCR amplification of *Ld*RK: lane 1, 100bp DNA ladder; lane 2, PCR product of RK. (B) Release of insert from pJET1.2 vector: lane 1, 1Kb DNA ladder; lane 2, Restriction digestion of pJET-*Ld*RK with *Nde*I and *Sal*I; lane 3, Uncut vector of pJET-*Ld*RK. (C) Restriction digestion of pET-28a-*Ld*RK: lane 1, 1kb DNA ladder; lane 2, Uncut construct of pET-28a -*Ld*RK.

3.1.4 Purification of recombinant *Ld*RK

Recombinant *Ld*RK was expressed in *E. coli* BL21 (DE3) and then purified using Ni^{2+}-affinity and gel-filtration chromatography with a final yield of 15 mg per litre of culture medium. Analysis of the pooled purified fraction by SDS-PAGE indicated that the protein was >95% pure with an approximate molecular weight of 38 kDa (Fig. 7A-B), which corresponds to the theoretical molecular mass (37.5 kDa) predicted by *ProtParam* [163] . Gel filtration profile of *Ld*RK indicates that the recombinant protein existed as a dimer in the solution (Fig. 8A-B).

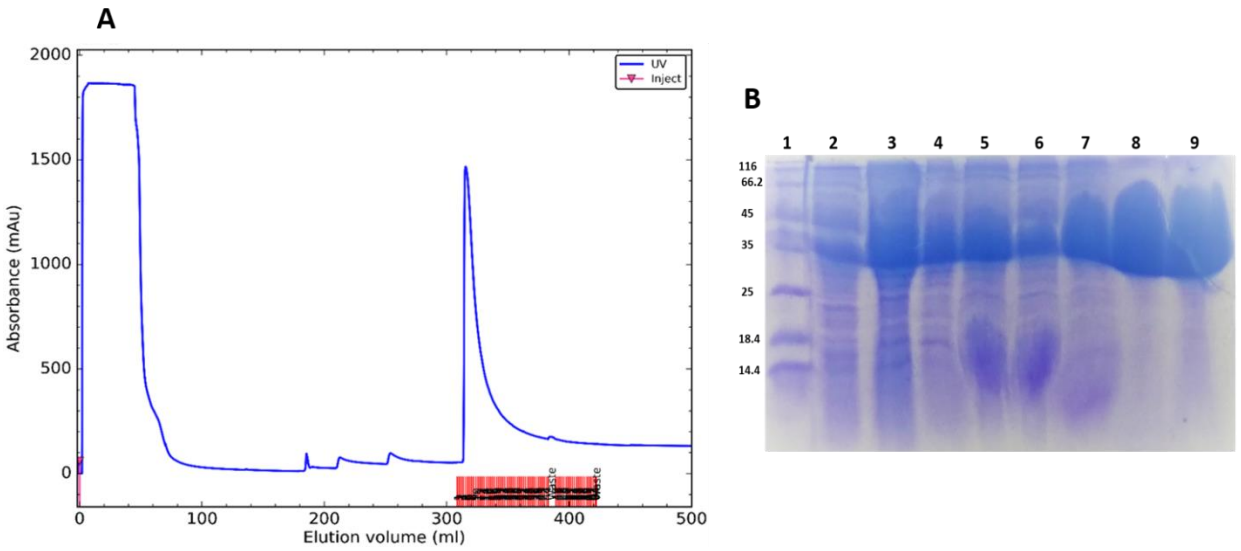


Figure 7. Purification of *Ld*RK by Ni^{2+}-NTA affinity chromatography. (A) Chromatogram depicting the elution profile of *Ld*RK from HisTrap HP 5ml column with a gradient of imidazole. (B) Coomassie blue stained 12% SDS-PAGE of RK purification profile: Lane 1, Molecular-weight markers (kDa) are indicated on the left size of gel; 2, Un-induced; 3, Induced; 4, Pellet; 5, Supernatant; 6, Flow through; 7-8, Peak fractions eluted with 400mM imidazole.

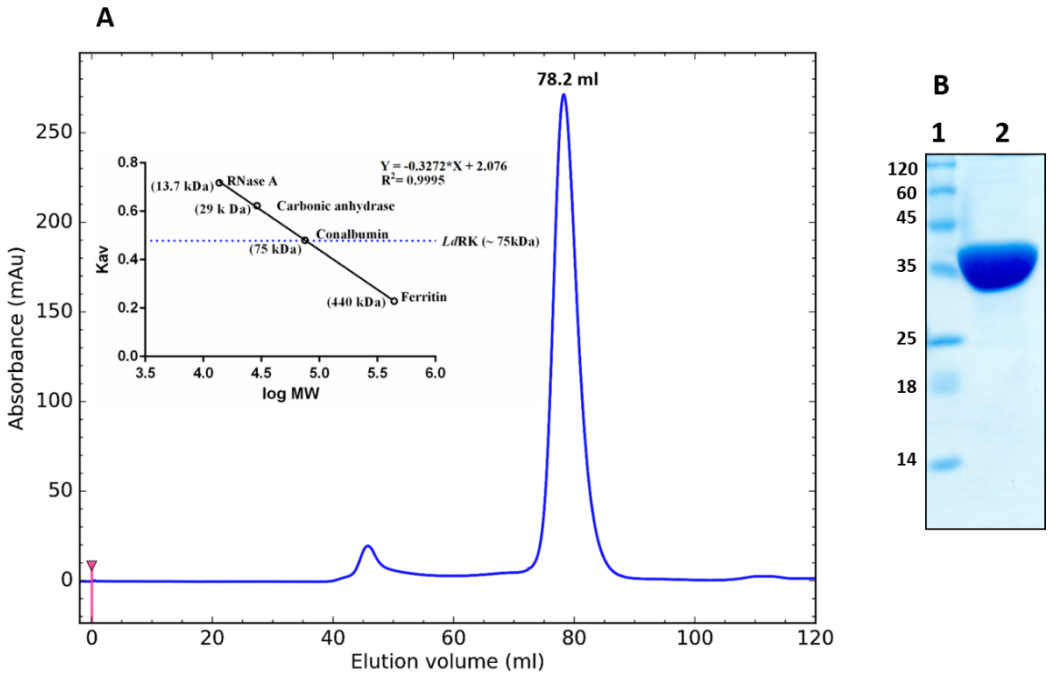


Figure 8. Purification of *Ld*RK by gel filtration chromatography. (A) Gel filtration profile from HiLoad 16/60 Superdex 200 pg column indicates the dimeric state of *Ld*RK. (B) 12% SDS–PAGE of pooled purified fractions after gel filtration chromatography. Lane 1, Molecular-weight markers (kDa) are indicated on the left size of gel; 2, Pooled purified fractions.

3.3 Discussion

In the current study, preliminary sequence analysis, cloning, expression and purification recombinant *Ld*RK was carried out. *Ld*RK constitutes 329 amino acids in the primary structure with two signature sequences, -NEVE- (209-212) and -GAGD- (275-278) of proteins belongs to the PfkB family. Comparative sequence analysis specified the high conservation of ribose and ATP binding motifs with RK sequences of *E. coli, S. aureus, V. cholerae* and *H. sapiens*. However, RK of *L. donovani* was differed from other RKs due to the two amino acids, F25 and N302 that are involved in the binding of ribose and ATP, respectively. Insertion of a five amino acid stretch at two the positions, 143-148 (-NYERI-) and 190-194 (-KPAEV-) is another unique feature observed in the protein sequence of *Ld*RK. Predicted crystallizability of full length construct (329 residues) was sub-optimal. The ORF coding for RK from *L. donovani* (*Ld*RK) was cloned in pET-28a expression vector and then recombinant protein was purified from *E. coli* BL21 (DE3) cells with a final yield of 15 mg per litre of culture medium. The purified recombinant ribokinase existed as dimer in solution, a frequently observed oligomeric sate in the members of RK family.

Chapter II

4. Biophysical and biochemical characterization of recombinant *Ld*RK

4.1 Results

4.1.1 Structural and stability studies of *Ld*RK by Circular dichroism spectroscopy

The secondary structure content of purified protein was analyzed by far-UV CD spectroscopy. The spectrum specifies α-helix as the dominant secondary structure in *Ld*RK due to the presence of characteristic double minimum near 222 nm and 208 nm, and a maximum around 192 nm (Fig. 9A). Analysis of the data with *DichroWeb* server [167] predicted the secondary structure content as follows: 27% of α-helix, 23% of β-sheet, 22% turns and 28% of unordered region. The presence of proportionately more α-helix in *Ld*RK correlates with crystal structures of ribokinase reported from *V. cholerae* [150], *S. aureus* [149] and *E. coli* [148]. To probe the thermal stability of ribokinase, change in the ellipticity was measured at 222 nm as a function of temperature from 293 to 343 K. The thermal melting (T_m) curve of ribokinase indicates a rapid increase in the ellipticity at 313 K which begins to level off at 323 K (Fig. 9B). The T_m of protein was calculated as 317.2 K by plotting the first derivative of thermal melting curve (Fig. 9C). Whereas, the apparent T_m of RK from *E. coli* was reported as 329 K though the RKs from both organisms exhibit similar oligomeric state in solution [168]. These results allowed us to speculate that the low sequence identity between two RKs may lead to the variation in thermal stability.

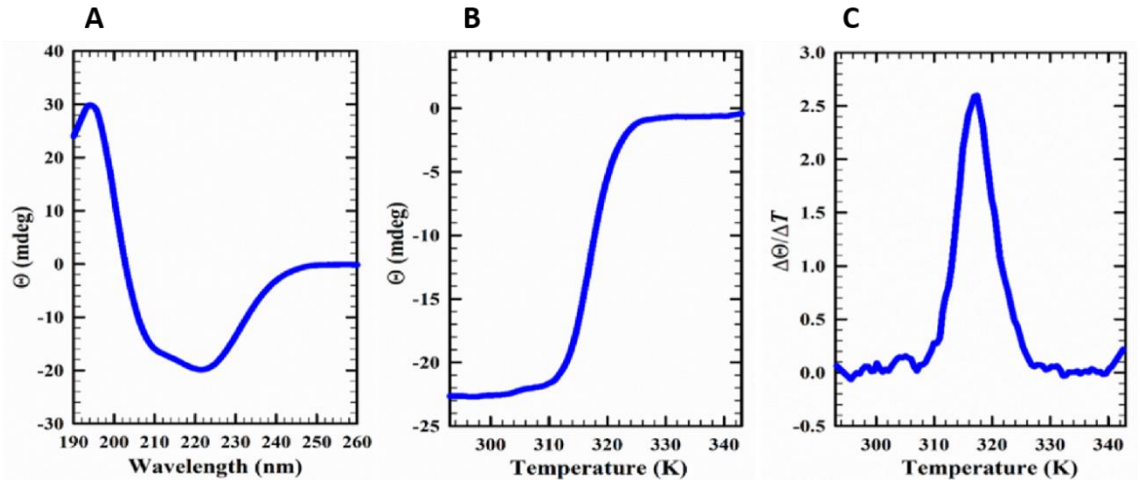


Figure 9. Structural and stability analysis of *Ld*RK by CD spectroscopy. (A) Far-UV Circular-dichroism spectrum of *Ld*RK measured at 298K in 7.5 m*M* Tris-HCl pH 8.5. (B) CD thermal denaturation curve of *Ld*RK obtained by measuring the ellipticity at 222 nm in 1.0 K intervals from 293 to 343 K. (C) The first derivative curve of thermal melting with peak value (*T*m) 317.2 K.

4.1.2 Functional studies of *Ld*RK

4.1.2.1 Enzyme kinetics

The kinetic parameters (K_m, k_{cat}, k_{cat}/K_m) were obtained by plotting the initial velocities versus respective substrate concentrations (Fig. 10A-B) and K_m values indicates that *Ld*RK has higher affinity for ATP than ribose (Table 3). *Ld*RK binds to ribose with the K_m of 296 µ*M*, which corroborates with the binding affinities (K_m) determined for RK of *L. major* (300 µ*M*) and human (279 µ*M*). Affinity of *Ld*RK towards ATP was observed to be almost similar (116 µM) to that of RK from *L. major* (200 µ*M*) [152] and human (70 µ*M*) [164].

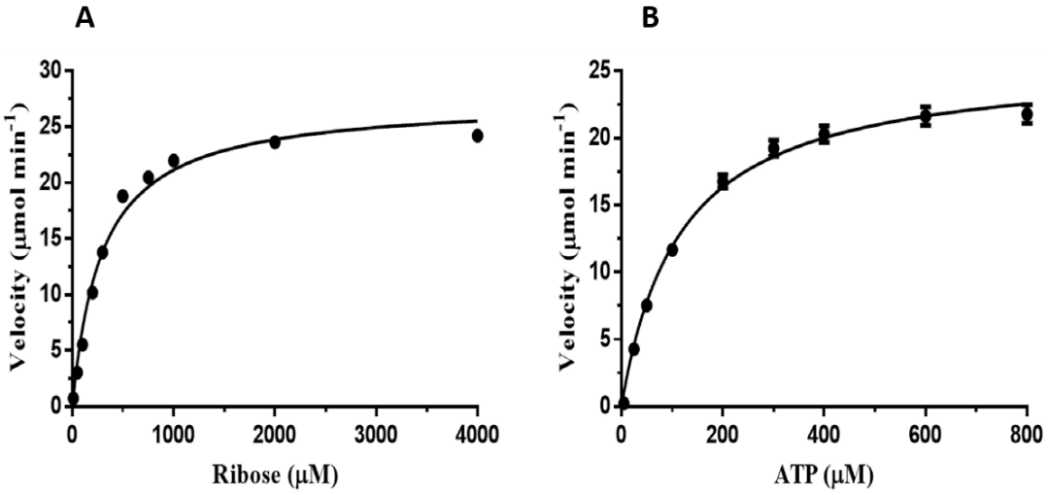


Figure 10. Kinetic characterization of *Ld*RK. (A) Plot of initial velocities against increasing concentrations of ribose. (B) Plot of initial velocities against increasing concentrations of ATP. Each data point represent average of assays performed in triplicate and standard error shown as error bars.

Table 3: Kinetic parameters of *Ld*RK

Substrate	**k_{cat} (s^{-1})**	**K_m (µ*M*)**	**k_{cat}/K_m (µM^{-1} s^{-1})**
Ribose	2.28 ± 0.08	296 ± 36	(7.7 ± 0.9) x 10^{-3}
ATP	2.15 ± 0.05	116 ± 9	(18.5 ± 1.5) x 10^{-3}

4.1.2.2 Regulation of RK activity by monovalent and divalent ions

Among the monovalent cations, K^+ has shown highest activating effect on *Ld*RK, which is similar to other known RKs (Fig. 11A). The activity got dropped to 90% and 45% with Cs^+ and Na^+, respectively, in comparison to K^+. Similar to *Arabidopsis thaliana* RK, the activity of *Ld*RK is could not be modulated by the pentavalent phosphate ion [136]. Due to the nonessential activation, *Ld*RK has shown some basal level activity in control reaction which has also been reported for human RK [165]. Most of the kinases, including the members of ribokinase family require Mg^{2+} as a metal cofactor for the hydrolysis of ATP. When magnesium was substituted with cobalt and manganese in the reaction mixture, the RK activity got reduced to 60% and 45% as compared to that with Mg^{2+} (Fig. 11B). When other divalent cations such as Ca^{2+} and Ni^{2+} were used, the activity dropped to as low as 5% and the activity got completely abolished with EDTA.

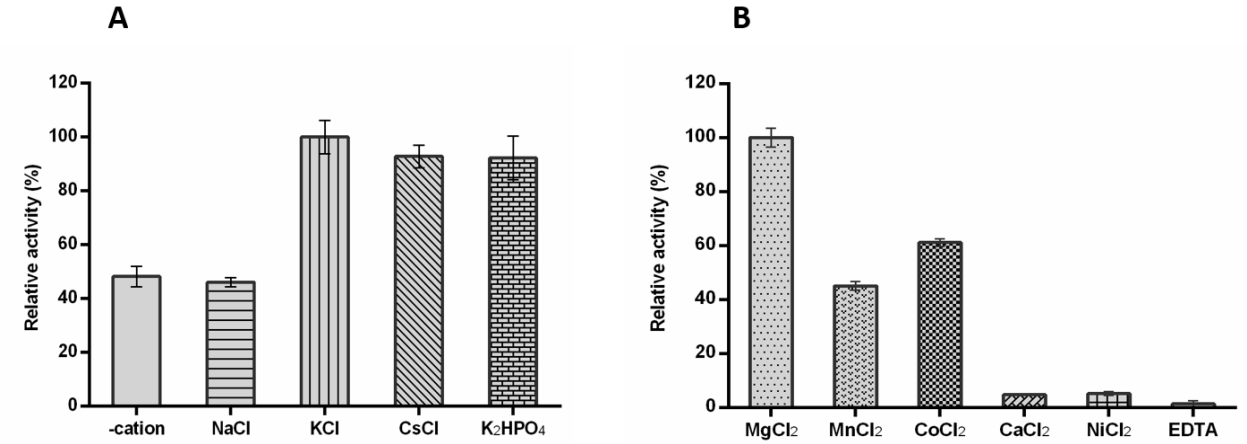


Figure 11. Dependence of *Ld*RK activity on monovalent and divalent cations. (A) For divalent cation assay, the concentration of divalent cations (Mg^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, Ni^{2+}) and EDTA was fixed at 1.5 mM and normalized to Mg^{2+}. (B) Activation of RK by monovalent cations. The concentration of each monovalent cation (Na^+, K^+ , Cs^+) and K_2HPO_4 was fixed at 100 mM and normalized to that of K^+. Each data point represent average of assays performed in triplicate and standard error shown as error bars.

4.2 Discussion

The biophysical and biochemical studies were used subsequently to characterize the purified recombinant RK. CD-spectroscopy results revealed the secondary structure information and thermal stability of purified recombinant protein. *Ld*RK has showed comparatively more α-helix in secondary structure content, a similar feature observed in the crystal structure of other RKs [148-150]. The thermal unfolding studies revealed that the T_m to be 317.2 K. Kinetic parameters were obtained by functional characterization of *Ld*RK, and K_m values for ribose and ATP was found to be 296 ± 36 µ*M* and 116 ± 9.0 µ*M*, respectively. Among the monovalent and divalent cations tested in biochemical studies, *Ld*RK has shown higher preference for Mg^{2+} and K^{+}, respectively. Though human and *E. coli* RKs have been reported to get activated by the presence of phosphate [165, 166], biochemical data suggest that *Ld*RK does not get activated by the presence of phosphate ions, similar to *At*RK as has been reported previously [136].

.

Chapter III

5. Crystallization and comparative structural analysis of apo and ligand bound structures of *Ld*RK

5.1 Results

5.1.1 Crystallization of *Ld*RK

5.1.1.1 Screening and crystallization of *Ld*RK-apo

Even though more than 950 conditions were screened to identify the preliminary crystallization conditions, crystals of *Ld*RK were obtained only with Index A7 containing 0.1 *M* Citric acid pH 3.5 and 3.0 *M* NaCl. Since crystals obtained through screen were not suitable for diffraction studies, condition was optimized further by varying protein concentration, pH and NaCl concentration of the precipitant solution. After refining the initial crystallization condition, long hexagonal rod-shaped crystals were obtained with 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl and 4.5% glycerol which reached to a maximum size after three weeks (Fig. 12A). The presence of a protein band at the corresponding molecular weight (~38 kDa) on silver stained gel confirmed the presence of *Ld*RK in crystals (Fig. 12B). Subsequently addition of glycerol in the above mentioned condition has tremendously improved size of the crystal.

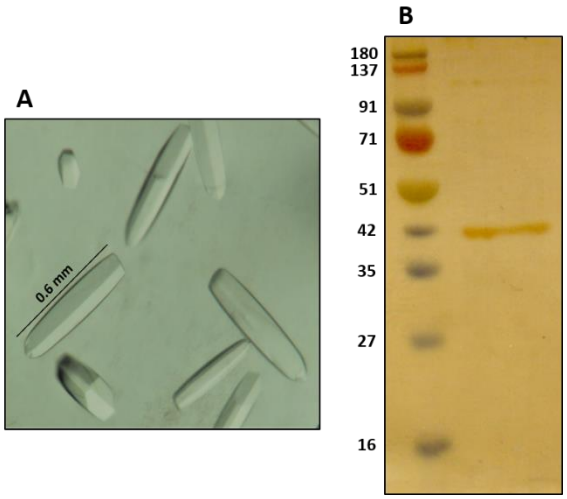


Figure 12. Crystallization of *Ld*RK-apo. (A) Crystals obtained with 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl and 4.5% (*v*/*v*) glycerol using hanging-drop vapor-diffusion method. (B) Silver stained SDS-PAGE of crystals.

5.1.1.2 Co-crystallization of *Ld*RK in complex with ATP, AMPCP and ADP

The optimized crystallization condition of apo protein was modified further to obtain co-crystals with ATP and its non-hydrolyzable analogue AMPPCP. When *Ld*RK was co-crystallized with ADP using the condition optimized for ATP or AMPCPCP, the rate of nucleation increased dramatically. Crystallization with the reduced protein concentration or varying the parameters of condition did not reduce the nucleation rate. Initial attempts with either controlled vapour diffusion using silicon oil or seeding allowed us to improve the crystal size (Fig. 13). When the seeding and controlled vapour diffusion methods were used simultaneously, the size and quality of the RK-ADP complex crystals enhanced than those obtained with either of the methods separately. The apo, ATP, ADP and AMPPCP bound crystals of *Ld*RK were obtained in space group $P6_1$ and diffracted to 1.95 Å, 1.8 Å, 2.01 Å and 1.98 Å resolutions, respectively (Fig. 14A-C). Two molecules of *Ld*RK were observed in the asymmetric unit for all the crystal structures reported here.

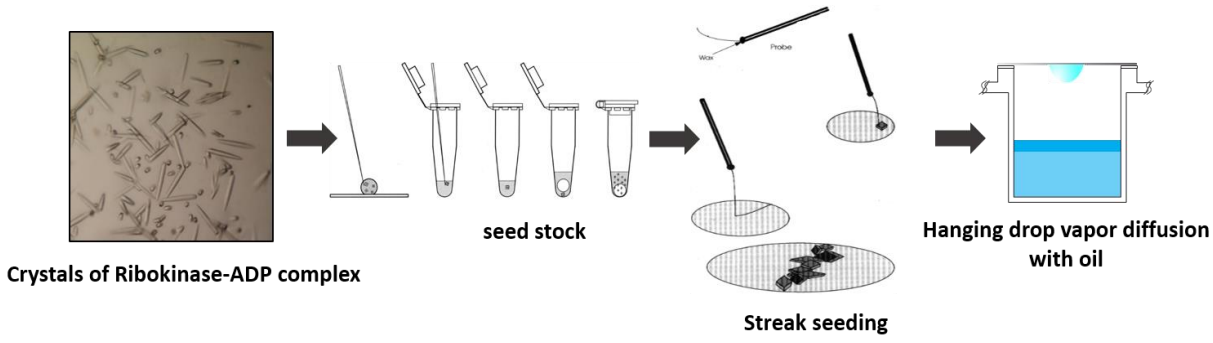


Figure 13. Crystallization of *Ld*RK-ADP. The crystals obtained in complex with ADP were used for preparation of seed stock. The seed was streaked over the drop and incubated in hanging position with oil barrier over reservoir solution.

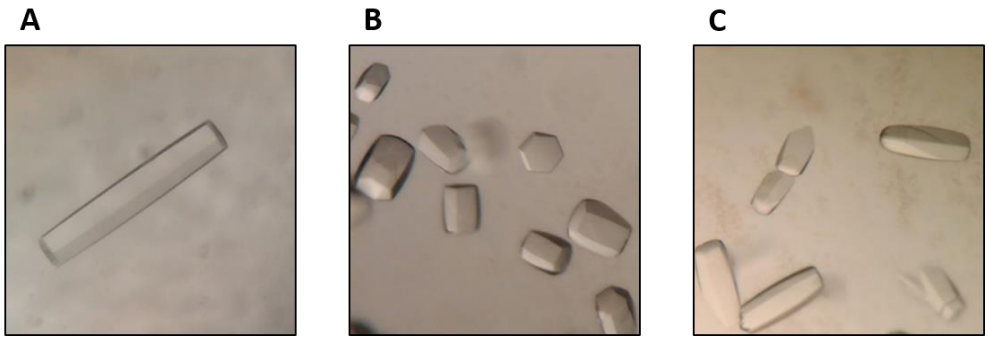


Figure 14. Crystals of *Ld*RK in complex with nucleotides. (A) *Ld*RK-ATP (B) *Ld*RK-AMPPCP (C) *Ld*RK-ADP.

5.1.2 Overall structure

To elucidate the structure-function relationship of ribokinase from a trypanosomatid parasite, we have determined the apo structure of RK from *L. donovani*. Each monomer constitutes 329 amino acids and both chains started at the third residue. The three dimensional structure of monomeric *Ld*RK consists of two distinct domains; a larger catalytic α/β domain and a small lid domain (Fig. 15). The α/β domain consists of three 3_{10} helices and a central, twisted eight-stranded β-sheet (β6/β5/β1/β9/β10/β11/β12/β13) packed against six α-helices (α4/α5/α6/α7/α8/α9) on one side and four α-helices (α1/α2/α11/α12) on the other side. The catalytic domain also contains a large ATP loop connecting α10 and β12, which plays a critical role in the binding of ATP and the overall topology resembles extended Rossmann fold. The lid domain is constituted by two pairs of antiparallel β-strands (β2-β4 and β6-β7) and a short β-strand (β3), connecting β2 and β4, along with a loop. The two domains are connected to each other by four polypeptide segments.

Similar to the RKs reported from *E. coli* (PDB: 1RKA) [159] and *V. cholerae* (PDB: 4XDA) [150], *Ld*RK also exist as a dimer in solution . In the crystal structure, the two protomers observed in the asymmetric unit are related by a non-crystallographic symmetry (Fig. 16). The dimerization of RK is facilitated by the orthogonal packing of five β-sheets around a two-fold axis that results in the formation of a flattened β barrel, also known as β-clasp. Analysis of the dimeric interface by PISA specified the buried surface area between two *Ld*RK molecules to be approximately 2636 $Å^2$ and ΔG^{disso} of 15.7 kcal/mol. A plentiful of interactions between the lid domains stabilizes the β-clasp and promote the dimerization. A total of 39 interfacing residues from each monomer interact through eighteen hydrogen bonds and one salt bridge. This salt bridge is formed at the interface between the D80 of chain A and K49 of chain B. Besides, the dimer was mostly stabilized by hydrophobic interactions formed by the residues V32, M35, M42, F47, L108, M110, L112, V123 and C125 present in the inside of the clasp.

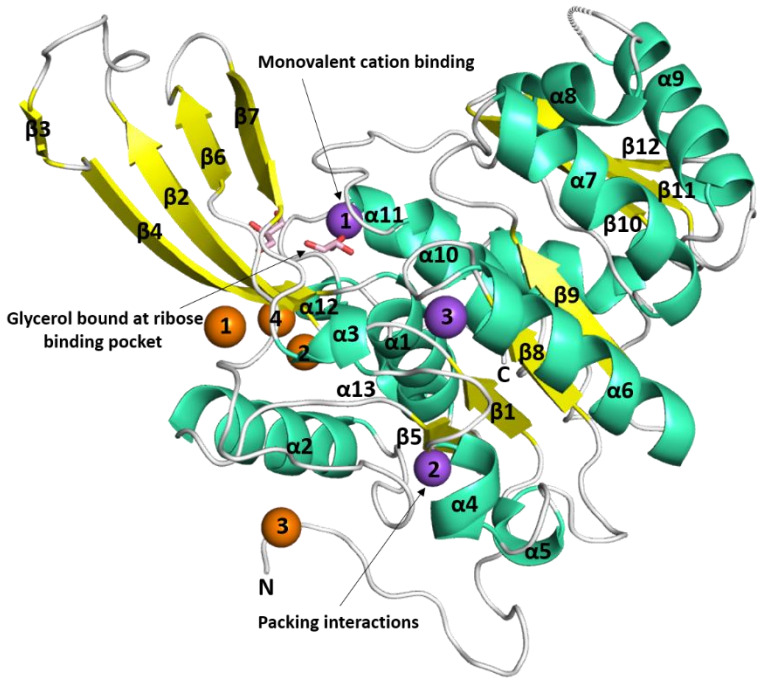


Figure 15. Schematic representation of monomeric subunit. The α-helices and β-sheets are colored cyan and yellow, respectively. Glycerol molecules bound to the monomer are show in pink sticks, whereas the sodium and chloride ions are indicated with violet and orange spheres, respectively.

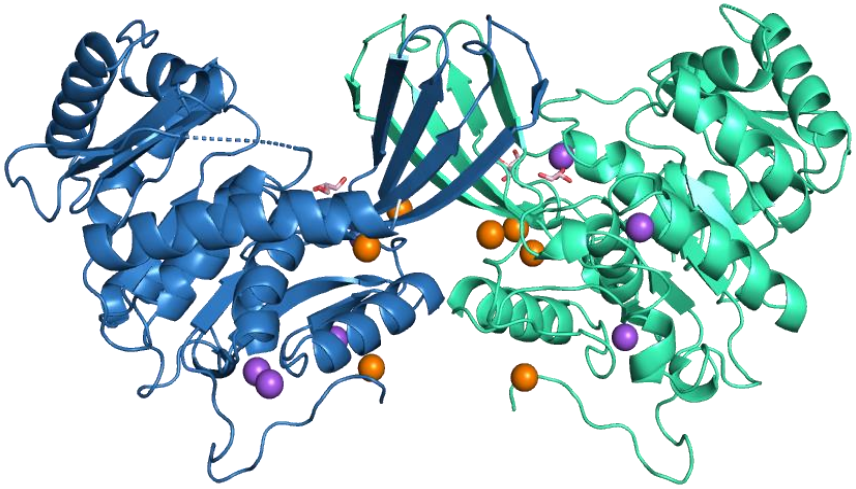


Figure 16. Dimeric structure of *Ld*RK observed in the asymmetric unit of crystal lattice. The subunit A and B are shown in blue and cyan, respectively.

Insertion of a five amino acid stretch at two positions in the primary structure makes *Ld*RK structurally different from other RKs reported so far (Fig. 17). The first insertion (143-147) that forms a 3_{10} helix, along with the loops on either ends resembles a 'loop-3_{10} helix-loop' architecture, while all other RK structures contain a bent helix in this region. The second insertion (189-193) leads to the formation of an extended α-helix (α7) in the corresponding region, compared to the shorter helix of other RKs. A structural homologue search using DALI sever indicated that *Ld*RK has the highest structural similarity to ribokinase from *H. sapiens* (PDB: 2FV7, Z-score – 42.3; r.m.s.d. – 2.1Å for 307 C_α atoms), *C. neoformans* (PDB: 6CW5, Z-score – 36.5; r.m.s.d. – 2.7Å for 290 C_α atoms) and adenosine kinase of *A. tumefaciens* (PDB: 2RBC, Z-score – 33.8; r.m.s.d. – 2.7Å for 293 C_α atoms). Three glycerol molecules were modeled in the structure of RK during the final rounds of refinement based on the difference density maps. Due to the disorder in loop regions, no electron density was observed for the residues in the stretches 266-272 and 317-329 in chain A and 221-222 in chain B.

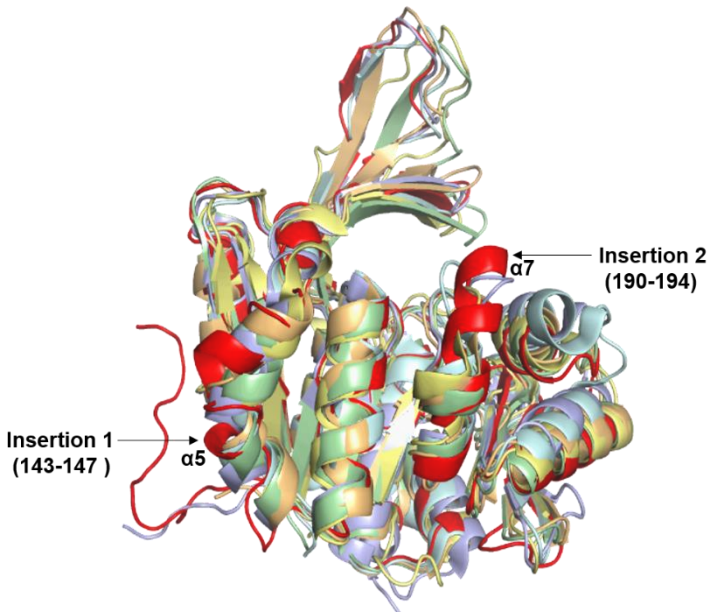


Figure 17. Structural alignment of *Ld*RK with other RKs. The first insertion (143-147) leads to the formation of helix-α5, whereas the second insertion (190-194) forms the extended helix-α7.

5.1.3 Binding of monovalent ions and activation of *Ld*RK

During refinement, difference map peaks were identified at various positions in the structure of RK. As 3.4 M NaCl was used in the crystallization condition, and based on coordination and bond lengths, Na^+ and Cl^- ions were assigned at multiple sites with full occupancy (Fig. 15). Among the multiple sodium-binding sites, binding of a Na^+ ion at the site-1 activates RK (discussed below), whereas at the site-2, it involves in crystal packing interactions. At site-2, one Na^+ ion occupies a symmetric position by binding to the backbone O atoms of A67 and G92 and two water molecules as well as $O^{\varepsilon 1}$, $O^{\varepsilon 2}$ of D135 from adjacent dimer that presents an octahedral geometry (Fig. 19A-B). Whereas, the Cl^- ion binds RK at several positions with a coordination ranging from 1-5 in each monomer (Fig. 20A-D). At site-1 a Cl^- ion was coordinated to backbone N atom of S105 and at site-2 it was found to coordinate to the side chain $N^{\eta 1}$ and $N^{\eta 2}$ atoms of R64 and a water molecule. At site-3, another Cl^- ion was tetra-coordinated to the backbone N of A4, N^{ε} and backbone N atoms of R5, and $N^{\eta 1}$ of R100. A fourth Cl^- site was observed to make penta-coordination with the backbone N atom of F51 and four water molecules. These results reveal the most frequent coordination of Cl^- ion with side chain and backbone N atoms of arginine, as has been reported previously [169].

Complexation of monovalent cations with enzymes facilitate the enzyme-substrate interaction and promote the catalysis in several ways [170]. The monovalent cations such as Na^+, K^+, Cs^+ and NH_4^+ are known to play a crucial role in the activation of RK [151, 165]. Binding of monovalent cations enhance the activity by organizing the nucleotide binding pocket of RKs [52, 149, 151, 154, 171]. Except the chain A of apo, all structures of RK are found to be activated by the binding of a Na^+ ion in the vicinity of ATP binding pocket. The coordination of Na^+ to the backbone O atoms of D272, R311, S317 and O^{γ} of S308 and two water molecules presents an octahedral geometry with Na^+- O distance range of 2.2-3.0 Å (Fig. 18A). The Na^+ ion fitted well in the density and got refined to an average B-factor of 22.4 $Å^2$ which is in agreement with the average B-factor 24.1 $Å^2$ of surrounding atoms. The superposition of *Ld*RK-apo onto the ribose-ADP bound structure of *Vc*RK (PDB: 4XDA) [150] revealed the identity of geometry and sodium-binding residues between these two structures (Fig. 18B).

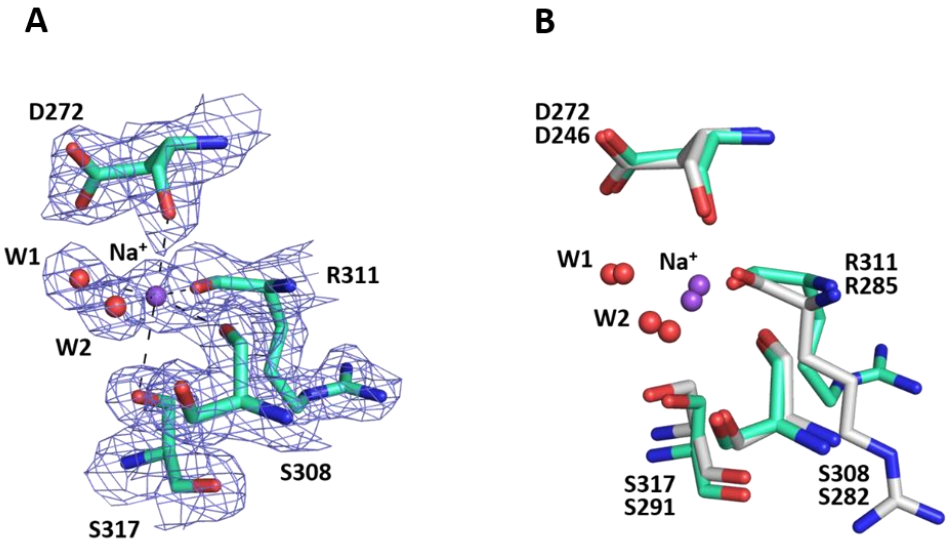


Figure 18. Binding of monovalent cation to *Ld*RK. (A) The $2f_o$-f_c map (blue mesh) of Na^+ binding residues contoured at 1.5σ. The Na^+ ion represented as a violet sphere is octahedrally coordinated by two water molecules along with the backbone O atoms of D272, S308, R311 and side chain O atom of S317. (B) Structural superposition of monovalent cation binding site from *Ld*RK-Na^+ and *Vc*RK-Na^+ (white; PDB: 4XDA). Interacting residues of the Na^+ binding site is represented in cyan (*Ld*RK) and white sticks (*Vc*RK-ribose-ADP, PDB: 4XDA).

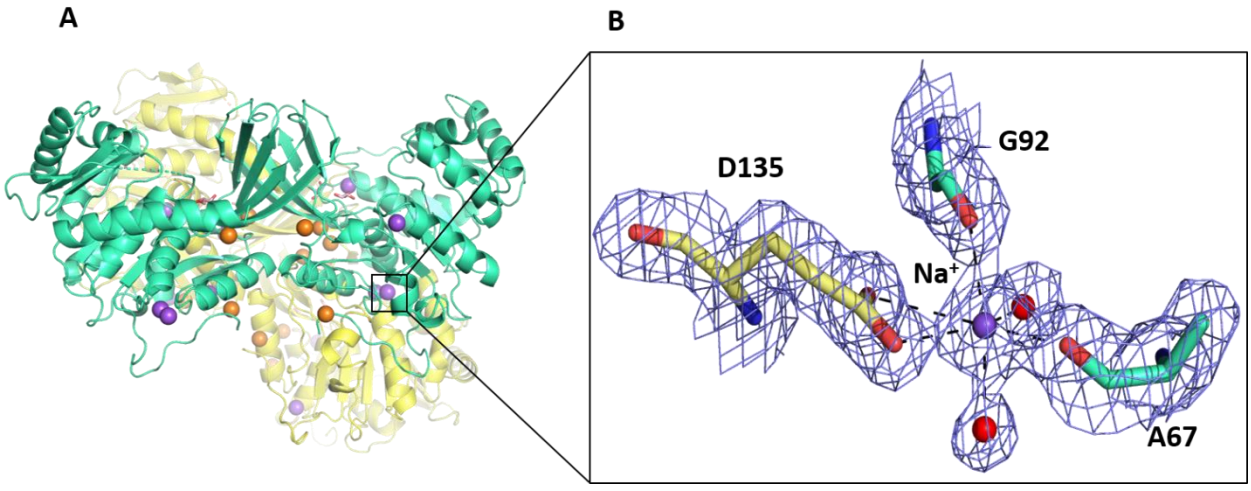


Figure 19. Role of sodium ions in crystal packing. (A) Crystal packing of *Ld*RK dimer with a Na^+ ion (site 2) at the interface. (B) A Na^+ ion makes lattice contacts by coordinating with the carbonyl O atoms of A67 and G92 along with two water molecules from one monomer and side chain O atoms of D135 from another adjacent dimer. The $2f_o$-f_c map (blue mesh) contoured at 1.0σ.

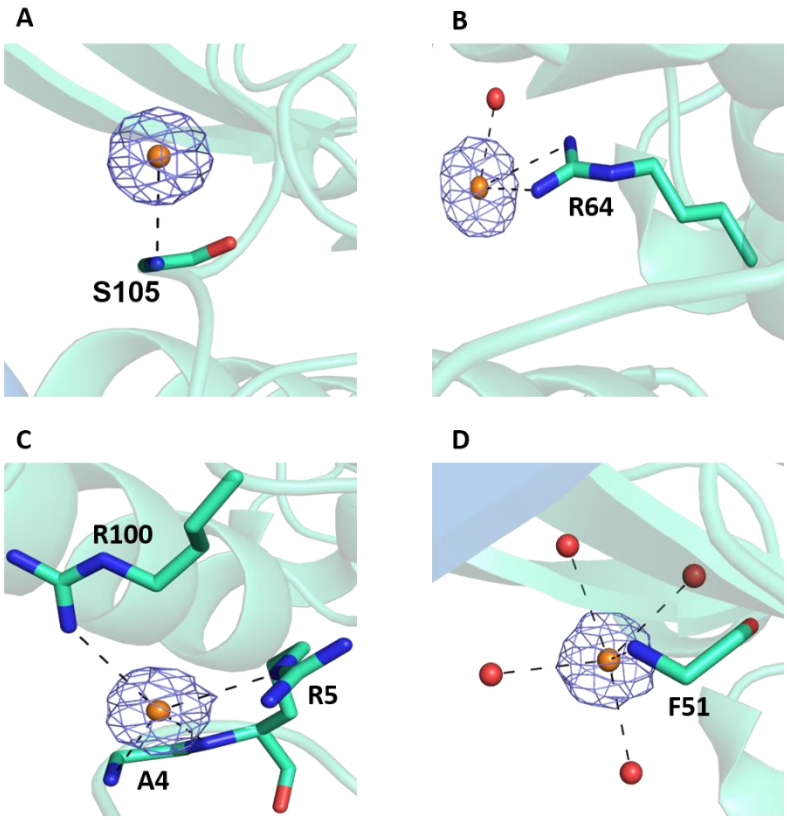


Figure 20. Binding of chloride ions to *Ld*RK. The chloride ions bound to RK at multiple sites with a coordination 1-5. Nitrogen was found to be the most frequent coordinating atom in all sites and Figure A-D represents Site 1-4, respectively. The $2f_o$-f_c map (blue mesh) contoured at 1.0σ.

5.1.4 Nucleotide-binding site

To decipher the mode of nucleotide binding and concomitant structural changes, the structures of *Ld*RK with ATP, ADP and AMPPCP were determined (Fig. 21A-C). The positive difference density (mF_o - DF_c, α_{calc}) map for the ligands was quite clear in the nucleotide binding pocket of their respective structures during the initial stages of refinement. In the subsequent description, we have used chain A from all complex structures unless stated otherwise.

Each monomer of *Ld*RK binds an ATP molecule in a shallow groove present at the C-terminal end of central β-sheet. The nitrogen base is held in an anti-conformation sandwiched between the side chains of A248 from short ATP loop and V309 from α10 by hydrophobic interactions (Fig. 22A). Besides, the adenine base is also stabilized by the solvent mediated hydrogen bonds; the N-1 atom interacting with the backbone N and O atoms of V268, while N-3 atom connecting to the backbone N atom of K266 and side chain N^{δ} of N302. The ribose ring exits in C3′- endo conformation, stabilized by the hydrophobic contacts with the phenyl ring of F280. In addition, the O2′ and O3′ atoms of ribose sugar forms direct hydrogen bonds with side chain N^{δ} of N302 and backbone O of G250, respectively. The N302 of *Ld*RK is structurally equivalent to H279 from *Ec*RK (PDB: 1RKA) [159] and H276 from *Vc*RK (PDB: 4XDA) [150].

The three phosphate groups of ATP are connected to the protein by direct and solvent-mediated hydrogen bonds (Fig. 22A). The O1A of α-phosphate accepts hydrogen bonds from the O^{γ} of T245 and backbone O of G247, while the O2A interacts with a water molecule. The O1B of β-phosphate coordinated to N^{δ} of N209, while a water molecule bridges the O2B with N^{δ} of N183 and $O^{\varepsilon 1}$, $O^{\varepsilon 2}$ of E212. The γ-phosphate group was anchored to the amide backbone of -GAGD- motif by an extended network of interactions. The O1G of γ-phosphate contacts the amide groups of A276 and G277 by direct hydrogen bonds, and with the G275 and D278 through intervening solvent molecules. The O1G makes an additional interaction with the backbone O atom of the residue T273, adjacent to the -GAGD- motif. The last residue (D278) of the motif acts as a catalytic base to abstract proton from the 5′-OH group, which is a prerequisite for nucleophilic attack on the γ-phosphate of ATP. The highly conserved -GAGD- motif is predicted to form an anion hole [148, 172] to stabilize the pentavalent state during the transfer of γ-phosphate group (transition state) from ATP to the 5′-OH of ribose.

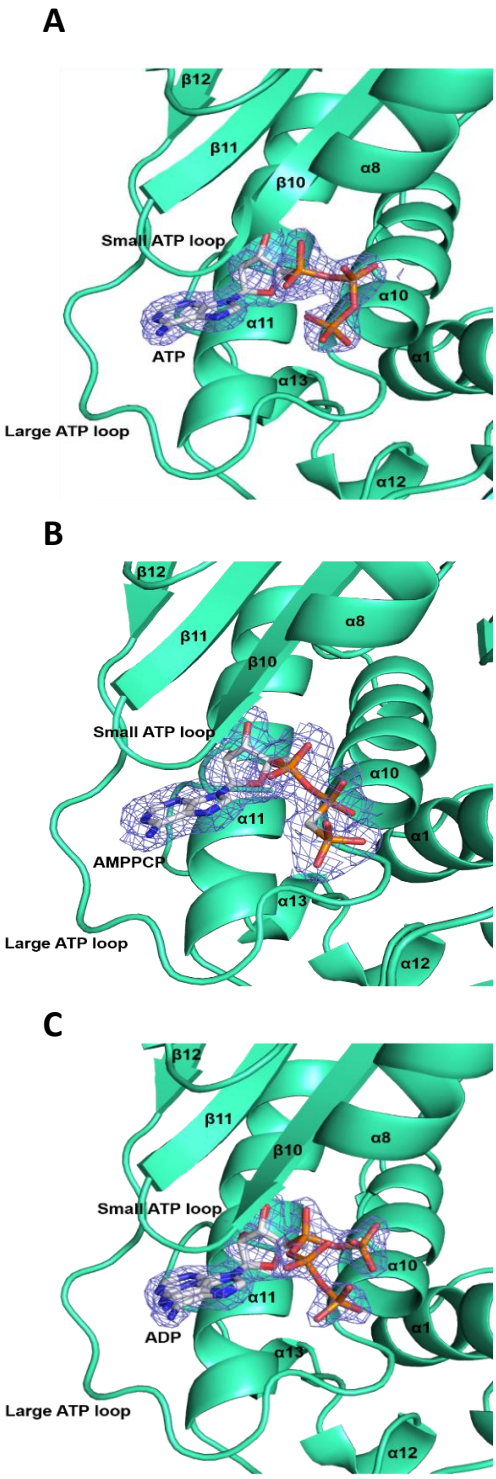


Figure 21. Nucleotide binding site of *Ld*RK. A, B, C denotes the $2f_o$-f_c map (blue mesh) of ATP, AMPPCP and ADP contoured at 1.0 σ in the subunit A of respective complex structures.

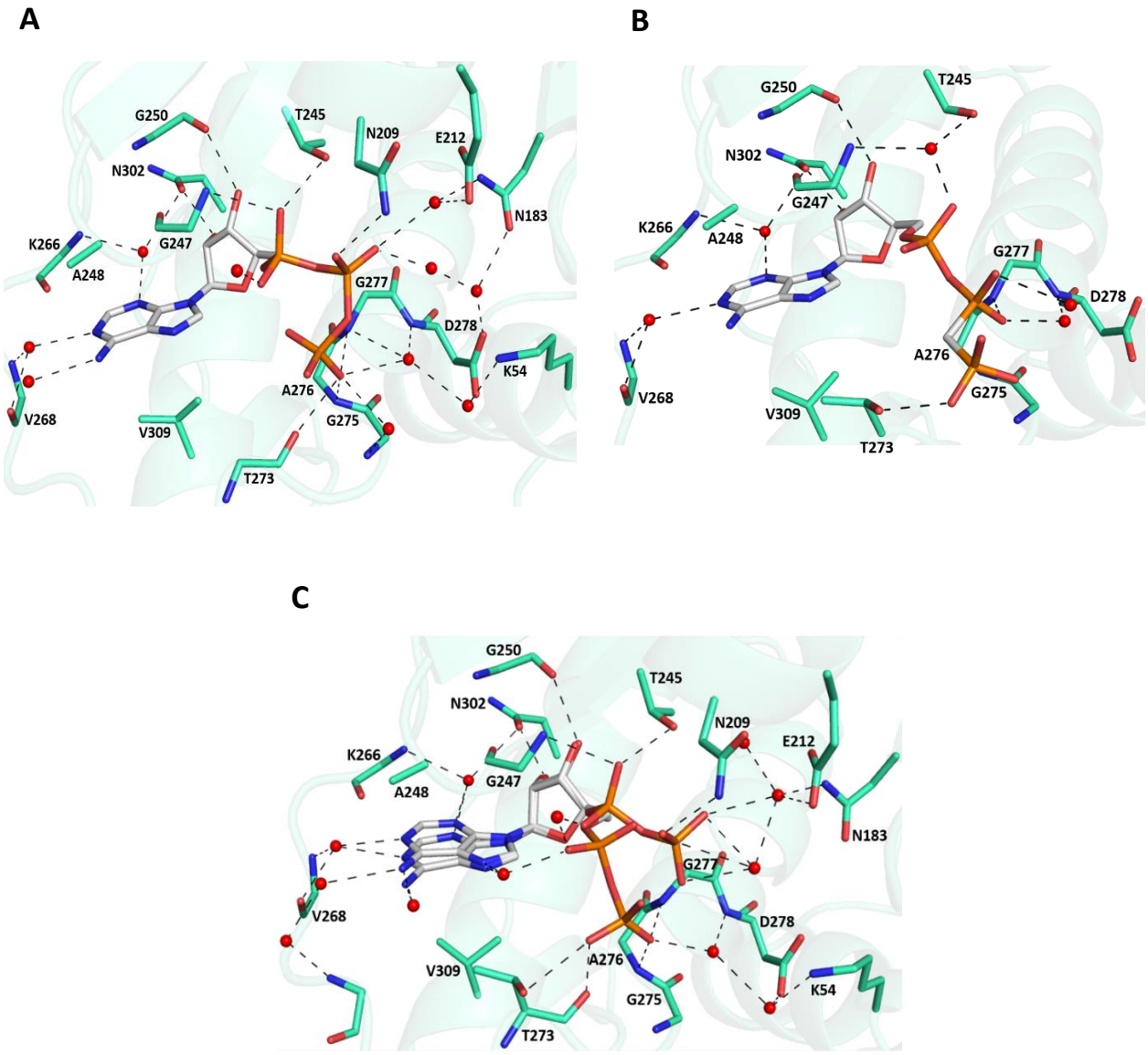


Figure 22. Interactions between *Ld*RK and nucleotides. (A) ATP contacted to residues in the nucleotide binding pocket by direct and water mediated hydrogen bonds. (B) Interaction of AMPPCP by less number of hydrogen bonds with a different conformation for β- and γ-phosphates. (C) ADP bound with an alternate conformation for β-phosphate, positioned similarly as γ-phosphate of ATP. Binding residues are shown as sticks and dashed lines represent the interatomic distances ≤ 3.2 Å.

Comparison of the AMPPCP binding to ATP in the *Ld*RK-ATP structure revealed major differences in the interactions between the enzyme and α-, β- and γ-phosphates. In particular, the substitution of β-γ linking oxygen (ATP) with carbon has significant consequences. Although the adenine and ribose moieties form similar environmental interactions as observed in ATP, the replacement by carbon in AMPPCP shifts the position of γ-phosphate away from the -GAGD- motif (Fig. 22B). The slight displacement of P^{α} atom by 0.69 Å leads to the replacement of direct hydrogen bonds between the α-phosphate and enzyme with solvent-mediated indirect hydrogen bonds. Surprisingly, the β-phosphate occupied the γ-phosphate position of ATP by linking O1B with the backbone N atoms of G277 and D278 (-GAGD- motif) by direct and indirect hydrogen bonding, respectively. As a result, the γ-phosphate was extended further away by making direct hydrogen bond from O1G to OG1 of T273. But we were unable to model an ambiguous density observed near the γ-phosphate neither as a ligand nor a solvent molecule.

The ADP molecule binding in the nucleotide binding pocket adopts an alternate conformation for the β-phosphate (Fig. 22C). In one possible conformation, the nitrogen base, ribose sugar, α- and β-phosphate groups of ADP forms environmental interactions aforementioned in RK-ATP. Interestingly, the position and interactions of β-phosphate in alternate conformation are similar to the γ-phosphate of ATP and an additional hydrogen bond formed between O^{γ} of T273 and O3B. Although 10 mM $MgCl_2$ was included in the crystallization buffer, no electron density was observed for Mg^{2+} ion between the phosphates of ATP, AMPPCP and ADP as well as at the conserved metal-binding motif –NXXE– [165].

5.1.5 Comparison of apo *Ld*RK structure with ATP, ADP and AMPPCP bound structures of *Ld*RK

To identify the ligand induced conformational changes, the structures of the *Ld*RK-ATP, *Ld*RK-AMPPCP and the *Ld*RK-ADP complexes were superposed onto the apo structure. All ligand bound structures resembled the open state conformation observed in the apo structure of RK (Fig. 23A). The overall folds of the RK with ATP, ADP and AMPPCP are very similar to that of apo structure with an r.m.s. deviation of 0.3 over 325 C_α atoms, indicating that these nucleotides binds to RK without inducing major conformational changes. The structures of *Ld*RK-apo and *Ld*RK-ATP, ADP and AMPPCP resemble the open state conformations of other ribokinases from *K. pneumoniae* (PDB: 3IKH) and *H. sapiens* (PDB: 5BYC). Except F280, the residues in the nucleotide binding pocket are relatively unaltered between the apo and the binary complex structures (Fig. 23C). Even though phenylalanine is not directly involved in the catalysis, this residue is conserved both in the sequence and tertiary structures of RKs reported so far (Fig. 4). The phenyl ring of F280 from α10 appeared 18° rotated to accommodate the ribose moiety of nucleotide by hydrophobic interactions (Fig. 23B). This observation corroborates with the 55° rotation of the phenyl ring of structurally equivalent F254 in the ADP-Ribose bound structure of *Vc*RK (4XDA) [150].

In addition, density for the region corresponds to missing residues of the large ATP loop in *Ld*RK-apo structure was appeared in all nucleotide bound structures (Fig. 24). The solvent mediated hydrogen bonds observed between the adenine base of nucleotides to K266 and V268 of large loop may resulted in the stability of the loop. In case of ADP complex structure an additional interaction was observed between base and A270 by a network of three water molecules.

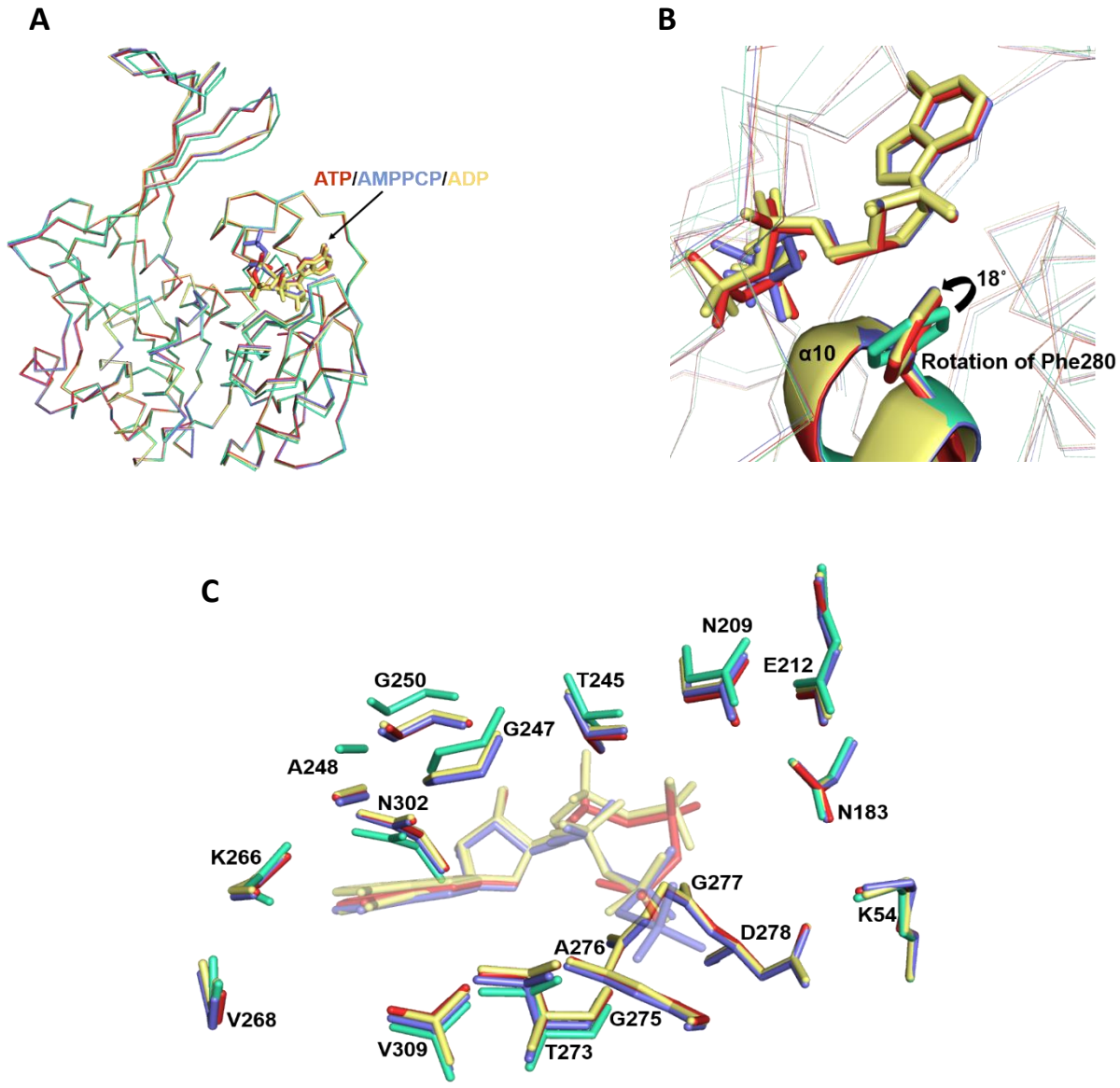


Figure 23. Comparison of apo and nucleotide bound structures of *Ld*RK. (A) Superposition of apo onto ATP, AMPPCP and ADP bound structures, which are represented as cyan, red, blue and yellow colored ribbons, respectively. (B) Rotation of F280 by 18˚ upon binding of nucleotide. (C) Diagram depicting structural superposition of ATP binding pocket of the *Ld*RK-apo onto *Ld*RK-ATP, *Ld*RK-AMPPCP and *Ld*RK-ADP.

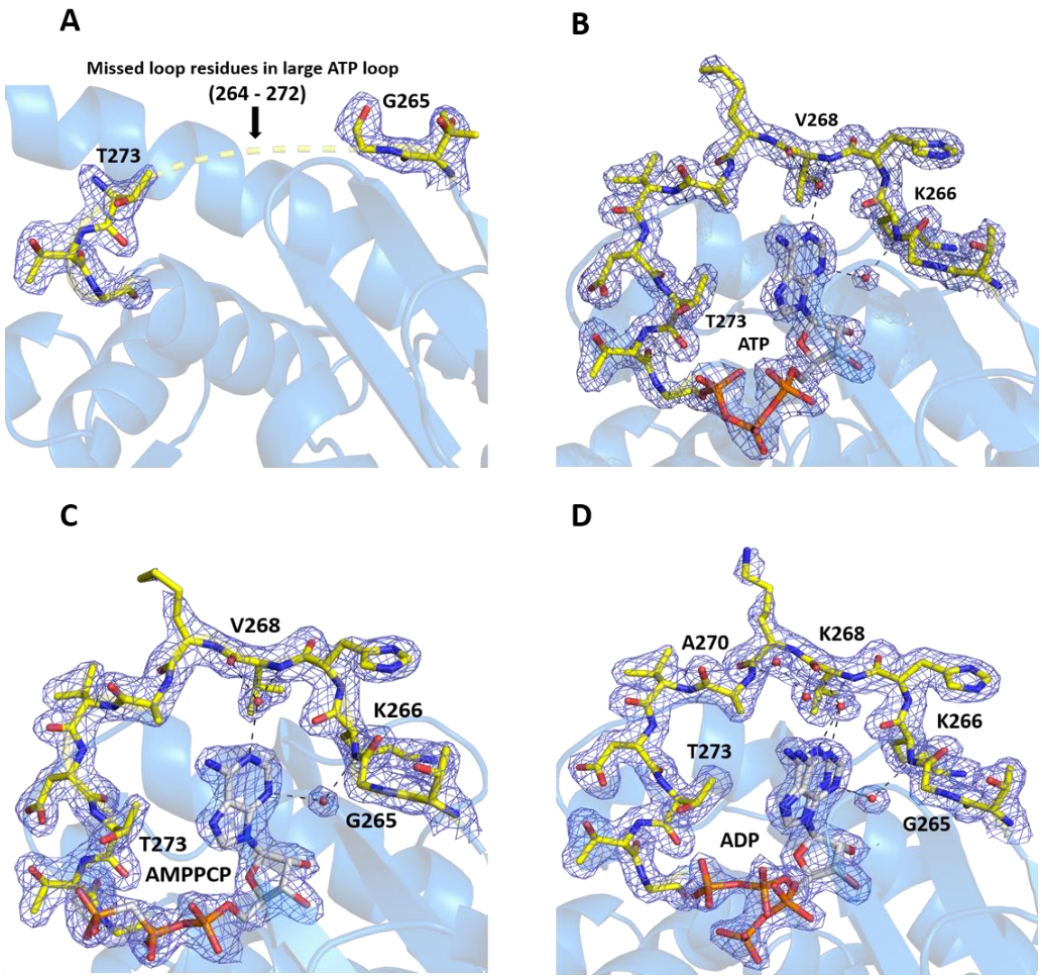


Figure 24. Stabilization of large ATP loop by solvent mediated interactions. (A) Density for part of the loop region corresponds to residues 264-272 were missed in the apo structure. (B)-(D) Appearance of density for large ATP loop in the ATP, AMPPCP and ADP bound structures due to solvent mediated intercations. The $2f_o$-f_c map (blue mesh) contoured at 1.0 σ.

5.1.6 Identification of substrate binding site and mechanism of *Ld*RK

Since attempts to obtain the ribose complex structure either by co-crystallization or by soaking failed, the ribose-binding pocket was identified by superposing the apo structure of *Ld*RK on to the ribose-ADP complex structure of *E. coli* (PDB: 1RKD) [159]. The structure of *Ld*RK was aligned to *Ec*RK (PDB: 1RKD) with an r.m.s.d. of 0.976 for 291 Cα atoms. The putative ribose-binding site of *Ld*RK is located in a deep cleft between the α/β domain and the lid domain. The residues forming direct hydrogen bonds with the oxygen atoms of ribose in *Ec*RK, N14, D16, G42, K43, N46, E143 and D255 are aligned well with F25, D27, G53, K54, N57, E160 and D278 of *Ld*RK, respectively (Fig. 25A). Consistent with the sequence alignment (Fig. 4), the structure alignment also revealed the conservation of ribose-binding residues between the *Ld*RK and *Ec*RK.

Among the three glycerol molecules modeled in the structure of RK, the glycerol molecule bound in the neighborhood of substrate binding pocket caught our attention and we did a close inspection of the interacting residues (Fig. 25B). The OH1 of glycerol is coordinated to D27, the OH2 forms hydrogen bonds to G53 and N57 through a water molecule, while the N57 forms hydrogen bond with the OH3 group of glycerol. These interactions with OH groups of glycerol partially mimic the substrate-protein interactions and may have prevented the accessibility of ribose into its binding pocket. Analogous to this, the crystal structures of several carbohydrate binding proteins such as lectins, glycoside hydrolases and glycosyltransferases have been reported with glycerol in the substrate-binding pocket [171, 173-176].

Even though the ternary structure of *Ld*RK with ATP and ribose was not obtained through crystallization, the relative orientation of γ-phosphate with respect to 5'-OH group of ribose can be anticipated by aligning the *Ld*RK-ATP structure modeled with the ribose-binding site. This ternary model structure was compared with the two complex structures of ATP available in PDB (Fig. 25C). The binding position and orientation of nucleotides in *Kp*RK (PDB: 3IKH) and *Ec*RKII (PDB: 3IQ0) are analogous to that of ATP molecule in the *Ld*RK. In all the three RK complex structures with nucleotides, the γ-phosphate appears to be projected towards the 5'-OH group of ribose, but in different orientations.

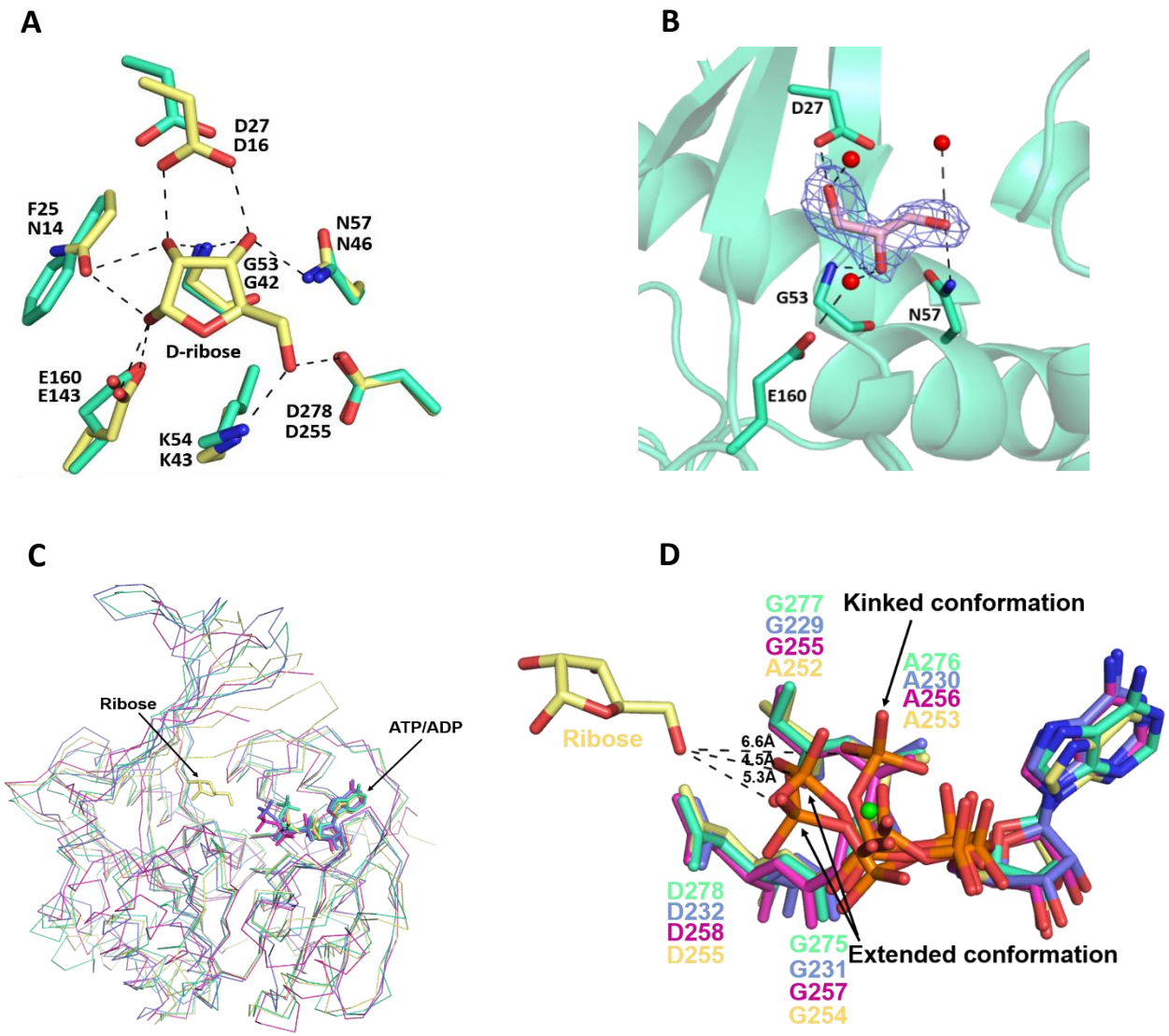


Figure 25. Identification of ribose binding pocket and mechanism of catalysis. (A) Ribose modeled into its binding pocket by the superposition of *Ld*RK-apo onto the *Ec*RK-ribose-ADP complex (PDB: 1RKD) structure. Ribose and its interacting residues from *Ec*RK are shown as yellows sticks, while the residues from *Ld*RK are shown as cyan sticks. Hydrogen bonds are highlighted with black colored dashes. (B) Hydrogen bonding interaction of glycerol with residues belongs the ribose binding pocket. The $2f_o$-f_c map for glycerol was contoured at 1σ. (C) Superposition of *Ld*RK-ATP (green cyan) onto *Ec*RK-ribose-ADP (PDB: 1RKD; yellow), *Kp*RK-ATP (PDB: 3IKH; blue) and *Ec*RKII-ATP-Mg^{2+} (PDB: 3IQ0; pink). (D) The distance between O5' and P^{γ} is represented by black dashed lines.

The γ-phosphate of *Ld*RK-ATP exists in a kinked conformation by pointing towards the adenine base, while it exists in an extended conformation in the case of *Kp*RK and *Ec*RKII (Fig. 25D). The oxygen atoms of the γ-phosphate of *Ld*RK-ATP interact with the backbone N-atoms of -GAGD- motif by an extensive network of direct and solvent-mediated hydrogen bonds. However, in ATP complex structures of *Kp*RK and *Ec*RKII, only an indirect hydrogen bond is formed between γ-phosphate and -GAGD- motif. The proposed catalytic mechanism for ribokinase family members indicates that the activated hydroxyl group on the substrate results in a nucleophilic attack on the γ-phosphate of ATP by an in-line displacement mechanism. The P^{γ} atom of ATP is located at a distance 6.6 Å from the O-5' atom of ribose in *Ld*RK-ATP, and 4.5 Å and 5.3 Å in *Kp*RK-ATP and *Ec*RKII-ATP (PDB: 3IQ0) complex structures, respectively, which are considerably higher than what is needed for the formation of a hydrogen bond for the phosphate group transfer between donor and acceptor.

5.2 Discussion

Ribokinase (RK) catalyzes the phosphorylation of ribose to produce ribose-5-phosphate in the presence of ATP and magnesium. *Ld*RK forms a dimer in the asymmetric unit with high structural similarity to RK structures reported from *E. coli* [159], *S. aureus* [149] and *V. cholerae* [150]. Structural superposition of *Ld*RK with other RKs revealed that the insertion of a five amino acid stretch observed in the sequence of *Ld*RK (section 3.1.1 of chapter 3) resulted in the formation of a short helix and an extended helix- α7.

Monovalent cation shapes the nucleotide binding pocket by changing the conformation of the large ATP loop to enhance the access for ATP into the binding pocket [151]. A Na^+ ion bound by two water molecules and four oxygen atoms from the protein presents an octahedral geometry, which is in contrast to the pentagonal bipyramidal geometry observed in case of *Ec*RK (PDB: 1GQT) [151] through the interaction of seven oxygen atoms from the protein. In support of the biochemical observations (section 4.3.2 of chapter 4), structural studies provided the basis for *Ld*RK to attain full activity upon binding to a monovalent cation. The crystallizability of *Ld*RK with only a single condition containing 3.4 *M* NaCl restricted our efforts to obtain the structure of RK in the inactive state.

The nucleotide binding pocket is located at one end of the α/β domain which is solvent accessible. The adenine base is anchored to the protein by hydrophobic contacts as well as solvent-mediated hydrogen bonds with residues from the large ATP loop connecting α9 and β10. This loop is indispensable for RK, as it comprises D272 that binds to Na^+, T274 which fixes the γ-phosphate of ATP and the region K266-V268 which helps to form the pocket that binds the adenine base of ATP in *Ld*RK. Lack of density for residues in the large ATP loop of chain A in apo structure indicates the conformational flexibility of this region. It was also reported that the large ATP loop attains variable conformations among the sixteen protomers present in the single asymmetric unit of *Vc*RK [150]. Surprisingly, improved density for the large ATP loop of chain A in nucleotide (ATP, AMPPCP and ADP) bound structures facilitated the modeling of loops in both chains of the dimer. Superposition of apo and ligand bound structures indicated that the large ATP loop attains the same conformation in all the structures. The ribose moiety of the nucleotide bound to the enzyme by direct hydrogen bonds through 2'- and 3'- OH groups and hydrophobic interactions is enabled by the rotation of phenylalanine side chain.

The phosphate moiety is connected by an extensive network of direct and indirect hydrogen bonds to protein. Though γ-phosphate is the actual functional group that gets transferred from ATP to 5′-OH of ribose during the catalysis, no RK structures have been reported so far to understand the interactions of γ-phosphate of ATP with RK. The intact density for the γ-phosphate in the current *Ld*RK-ATP complex structure allowed us to understand its interactions with the protein. The γ-phosphate of ATP is positioned at the conserved -GAGD- motif present at the N-terminal end of helix α8, which is thought to stabilize the penta-covalent transition state resulting upon the nucleophilic attack of O5' on the γ-phosphate of ATP [159]. An extensive network of hydrogen bonds, majorly the direct hydrogen bonds that strongly anchor the γ-phosphate to -GAGD- motif, resulted in a kinked conformation of ATP in *Ld*RK, whereas the lack of direct hydrogen bonds have led to an extended conformation of γ-phosphate in *Kp*RK-ATP complex structure (PDB: 3IKH). The flexibility of terminal phosphate group of ATP may give rise to differential projections as observed in both the structures.

Though ATP acts as a phosphate group donor, the metal cation-ATP complex is the true substrate for most of the kinases during catalysis. By binding to β-γ phosphates, the divalent cation not only acts as an electrophilic catalyst but also orients the γ-phosphate in-line with ribose for phosphoryl transfer. In this work, the acidic (pH 4.3) crystallization condition resulting in partial protonation of the phosphate groups could have hindered the binding of Mg^{2+} to ATP. In addition, the presence of 0.1 *M* citric acid in the crystallization condition could also have hampered the Mg^{2+}-ATP complex formation by reducing the effective magnesium concentration. Lack of divalent cations in the nucleotide bound structures states that the divalent cations though required for catalysis, but it may not be required for the binding of nucleotides to RK. The phosphate tail of Mg^{2+}-ATP complex structure of *Ec*RKII (PDB: 3IQ0) has an extended conformation with only a minor variation in the orientation of terminal phosphate. The variable distance between the O-5' and P^{γ} of the three ATP bound structures indicates that the structural changes accompanied by the binding of ribose may reduce the distance between the two substrates to facilitate phosphate group transfer. Comparative structural analysis indicates that both apo and nucleotide bound structures exist in an open state conformation. Lack of major structural rearrangements in ligand bound structures might be an adaptation to avoid the premature hydrolysis of ATP before the binding of ribose to RK.

This observation also specifies that the closed state of enzyme is not imperative for the binding of ATP to facilitate the accessibility of ribose to RK.

Except for the phosphate moiety, the mode of AMPPCP binding remains similar to that of ATP. Even though the phosphate moiety has modest interactions with the enzyme, the average B-factor of AMPPCP (38.80 Å) is almost identical to that of ATP (38.42 Å) and the overall structure of *Ld*RK-AMPPCP remains similar to that of RK-ATP structure (r.m.s.d. of 0.149 for 310 C_α atoms). Though AMPPCP is used as a stable analogue to study the interactions of ATP with the enzyme in several structures, the current study shows that AMPPCP may not be a suitable analog to study interactions of ATP with RK. An observation of the interactions between the γ-phosphate group of ATP or AMPPCP reveals that only one of the three non-bridging oxygens are connected to the protein by direct hydrogen bonds. The average B-factor of ADP (17.70 Å) is only half of the B-factor of ATP (38.42 Å) and AMPPCP (38.80 Å), whereas the B-factors for the respective protein chains are 31.18 Å, 34.71 Å and 21.74 Å. These results show that the enzyme gets better stabilized upon ADP binding.

We have not been successful in obtaining the structure of ribose bound form of *Ld*RK through crystallization. The ribose-binding residues predicted through structural superposition indicated that most of the residues in the binding pocket are indeed conserved. The substrate binding pocket of *Ec*RK is organized by seven residues that makes eleven direct hydrogen bonds with the four exocyclic oxygens of ribose [148]. Among the substrate binding residues, three are acidic (D16, D143 and D255), one is basic (K43), two are polar (N14 and N46) and one is non-polar (G42). Apart from G42, all residues contacted the -OH groups of ribose through their polar side chains. Interestingly, the polar N14 of *Ec*RK interacting with the O2' and O3' of ribose by two direct hydrogen bonds aligned with a non-polar F25 in *Ld*RK. The substituted phenylalanine may support the binding of ribose by stacking interactions in addition to the hydrogen bonds from other interacting residues.

Both subunits of apo and subunit B of nucleotide bound structures were bound to a glycerol molecule between the two domains of RK. The interactions with the -OH groups of glycerol are formed by more than half of the residues from the ribose-binding pocket. The interactions of glycerol with the ribose-binding residues mimics the enzyme-substrate based interactions and

might have hampered the binding of ribose. The presence of glycerol in the crystallization condition has led to the differential accessibility of ligand to active site as has been observed in the dimeric structure of *M. tuberculosis* RK-ribose complex (PDBs : 3GO6 and 3GO7), where one molecule in the asymmetric unit is bound to ribose and the respective site in the other subunit is occupied by glycerol.

6. Summary

D-ribose is a vital pentose sugar in all three kingdoms of life. It can be used as a component of several important biological macromolecules such as RNA, DNA and nucleotide co-factors as well as used as a source of energy. *Leishmania* is an auxotrophic protozoan parasite which acquires d-ribose by transporting it from the host cell and also by the hydrolysis of nucleosides. The enzyme ribokinase (RK) catalyzes the first step of ribose metabolism by phosphorylating ribose using ATP to produce ribose-5-phosphate. Though RKs from bacteria and human have been vastly explored at the level of molecular structure and function, the knowledge about RK from trypanosomatids is scarce. In the current study, we have attempted to understand catalytic mechanism of RK from *L. donovani* (*Ld*RK) by structural and functional approach.

The ORF encoding the full length RK from *L. donovani* was cloned, expressed and purified to homogeneity. Purified recombinant *Ld*RK exists as a dimer in the solution with comparatively more α-helix in the secondary-structure content. Kinetic parameters were obtained by functional characterization which revealed that *Ld*RK has higher affinity for ATP than ribose. The enzyme has shown preference towards the monovalent cation K^+ for activation along with strict requirement for divalent cation Mg^{2+}. Furthermore, the kinase activity could not be modulated in the presence of pentavalent phosphate ion.

The crystal structures of *Ld*RK apo and in complex with ATP, an ATP analogue AMPPCP and ADP were determined with a Na^+ ion present in monovalent cation binding site. Though structural studies facilitated the identification of monovalent cation binding site, biochemical studies specifies the preference of cation binding site of RK towards ions with radius greater than Na^+ for activation. Insertion of a five amino acid stretch at two positions forms helix that makes *Ld*RK structurally different from other RKs. Lack of complete electron density for large ATP loop in the chain A of apo structure denoted its conformational flexibility. The solvent mediated hydrogen bonds played a major role in the stabilization of large ATP loop upon complexation with nucleotides, which is evidenced by the appearance of electron density.

The γ-phosphate in *Ld*RK-ATP complex structure bound strongly to the conserved -GAGD- motif by an extensive network of direct and indirect hydrogen bonds, whereas the phosphate moieties of non-hydrolayzable ATP analogue AMPPCP makes different environmental interactions with RK due to the substitution of β-γ linking oxygen (ATP) with carbon .While, the co-product ADP bound

to RK with an alternate conformation for α and β-phosphates. Binding of ADP stabilizes the structure of RK compared to other two nucleotides (ATP and AMPPCP).

Lack of major structural changes between apo and nucleotide bound (ATP, AMPPCP and ADP) structures might be an adaptation to prevent the premature hydrolysis of ATP and the binding of nucleotide is not imperative for the ribose accessibility to RK. Binding of nucleotide (ATP, AMPPCP and ADP) to RK induced rotation of F280 to accommodate ribose moiety of nucleotide into its binding pocket by hydrophobic interactions. Comparative structural analysis with other ATP complex structures of RK indicates the flexibility of γ-phosphate leads to variable conformations of terminal phosphate group and the accompanied structural changes resulted by the binding of ribose are essential for phosphate group transfer. Presence of glycerol in the crystallization condition resulted in partial mimicking of substrate-enzyme interactions which hampered the accessibility of ribose to RK. The predicted substrate pocket of *Ld*RK differed with other RKs by the replacement of conserved polar asparagine with a non-polar phenylalanine.

Overall, these observations in general enhance the understanding of the catalytic mechanism of RK. In addition, RK is one of the possible sources for the generation of ribose-5-phosphate in *Leishmania* parasites. Hence the structure of RK can be used further to design an inhibitor which prevents the salvage of free ribose produced through degradation of nucleosides and its reuse in the biosynthesis of nucleotides.

7. References

1. Stuart, K., et al., *Kinetoplastids: related protozoan pathogens, different diseases.* J Clin Invest, 2008. **118**(4): p. 1301-10.
2. Hannaert, V., et al., *Evolution of energy metabolism and its compartmentation in Kinetoplastida.* Kinetoplastid Biol Dis, 2003. **2**(1): p. 11.
3. Kaufer, A., et al., *The evolution of trypanosomatid taxonomy.* Parasit Vectors, 2017. **10**(1): p. 287.
4. Donovan, C., *On the possibility of the occurrence of trypanosomiasis in India. 1903.* Natl Med J India, 1994. **7**(4): p. 196, 201-2.
5. Leishman, W.B., *On the possibility of the occurrence of trypanosomiasis in India. 1903.* Natl Med J India, 1994. **7**(4): p. 196-200.
6. Ross, R., *Note on the Bodies Recently Described by Leishman and Donovan.* Br Med J, 1903. **2**(2237): p. 1261-2.
7. Murray, H.W., et al., *Advances in leishmaniasis.* Lancet, 2005. **366**(9496): p. 1561-77.
8. Okwor, I. and J.E. Uzonna, *The immunology of Leishmania/HIV co-infection.* Immunol Res, 2013. **56**(1): p. 163-71.
9. Akhoundi, M., et al., *A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies.* PLoS Negl Trop Dis, 2016. **10**(3): p. e0004349.
10. David, C.V. and N. Craft, *Cutaneous and mucocutaneous leishmaniasis.* Dermatol Ther, 2009. **22**(6): p. 491-502.
11. Pace, D., *Leishmaniasis.* J Infect, 2014. **69 Suppl 1**: p. S10-8.
12. Zhang, W.W. and G. Matlashewski, *Characterization of the A2-A2rel gene cluster in Leishmania donovani: involvement of A2 in visceralization during infection.* Mol Microbiol, 2001. **39**(4): p. 935-48.
13. Engwerda, C.R., M. Ato, and P.M. Kaye, *Macrophages, pathology and parasite persistence in experimental visceral leishmaniasis.* Trends Parasitol, 2004. **20**(11): p. 524-30.
14. De Trez, C., et al., *iNOS-producing inflammatory dendritic cells constitute the major infected cell type during the chronic Leishmania major infection phase of C57BL/6 resistant mice.* PLoS Pathog, 2009. **5**(6): p. e1000494.
15. de Vries, H.J., S.H. Reedijk, and H.D. Schallig, *Cutaneous leishmaniasis: recent developments in diagnosis and management.* Am J Clin Dermatol, 2015. **16**(2): p. 99-109.
16. Burza, S., S.L. Croft, and M. Boelaert, *Leishmaniasis.* Lancet, 2018.
17. Peters, W. and R. Killick-Kendrick, *The Leishmaniases in biology and medicine*. 1987, London ; Orlando: Academic Press.
18. Jopling, W.H., *Long incubation period in kala-azar.* Br Med J, 1955. **2**(4946): p. 1013.
19. Marsden, P.D. and R.R. Nonata, *Mucocutaneous leishmaniasis: a review of clinical aspects.* Revista da Sociedade Brasileira de Medicina Tropical, 1975. **9**: p. 309-326.
20. Torres-Guerrero, E., et al., *Leishmaniasis: a review.* F1000Res, 2017. **6**: p. 750.
21. Weigle, K. and N.G. Saravia, *Natural history, clinical evolution, and the host-parasite interaction in New World cutaneous Leishmaniasis.* Clin Dermatol, 1996. **14**(5): p. 433-50.
22. Ahluwalia, S., et al., *Mucocutaneous leishmaniasis: an imported infection among travellers to central and South America.* BMJ, 2004. **329**(7470): p. 842-4.
23. Pedras, M.J., et al., *Mucosal leishmaniasis: the experience of a Brazilian referral center.* Rev Soc Bras Med Trop, 2018. **51**(3): p. 318-323.

24. Blum, J., et al., *Treatment of cutaneous leishmaniasis among travellers.* J Antimicrob Chemother, 2004. **53**(2): p. 158-66.
25. Konecny, P. and D.J. Stark, *An Australian case of New World cutaneous leishmaniasis.* Med J Aust, 2007. **186**(6): p. 315-7.
26. Bennis, I., et al., *Psychosocial impact of scars due to cutaneous leishmaniasis on high school students in Errachidia province, Morocco.* Infect Dis Poverty, 2017. **6**(1): p. 46.
27. Zijlstra, E.E., *PKDL and other dermal lesions in HIV co-infected patients with Leishmaniasis: review of clinical presentation in relation to immune responses.* PLoS Negl Trop Dis, 2014. **8**(11): p. e3258.
28. Marsden, P.D., *Mucosal leishmaniasis ("espundia" Escomel, 1911).* Trans R Soc Trop Med Hyg, 1986. **80**(6): p. 859-76.
29. Organization, W.H., *First WHO Report on Neglected Tropical Diseases: Working to overcome the global impact of neglected tropical diseases.(2010)*. 2018, Geneva: WHO Press Google Scholar.
30. Braga, A.S.d.C., A.C.d.C. Toledo Junior, and A. Rabello, *Factors of poor prognosis of visceral leishmaniasis among children under 12 years of age. A retrospective monocentric study in Belo Horizonte, State of Minas Gerais, Brazil, 2001-2005.* Revista da Sociedade Brasileira de Medicina Tropical, 2013. **46**: p. 55-59.
31. Ashkan, M.M. and K.M. Rahim, *Visceral leishmanisis in paediatrics: a study of 367 cases in southwest Iran.* Trop Doct, 2008. **38**(3): p. 186-8.
32. Grech, V., et al., *Visceral leishmaniasis in Malta--an 18 year paediatric, population based study.* Arch Dis Child, 2000. **82**(5): p. 381-5.
33. Zijlstra, E.E. and A.M. El-Hassan, *3. Visceral leishmaniasis.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2001. **95**: p. S27-S58.
34. Collin, S., et al., *Conflict and kala-azar: determinants of adverse outcomes of kala-azar among patients in southern Sudan.* Clin Infect Dis, 2004. **38**(5): p. 612-9.
35. Davies, C.R., et al., *Leishmaniasis: new approaches to disease control.* BMJ, 2003. **326**(7385): p. 377-82.
36. Okwor, I. and J. Uzonna, *Persistent parasites and immunologic memory in cutaneous leishmaniasis: implications for vaccine designs and vaccination strategies.* Immunol Res, 2008. **41**(2): p. 123-36.
37. EB, E., *Elevated cortisol level due to visceral leishmaniasis and skin hyper-pigmentation are causally related.* Int J Sci Commer Humanit, 2014. **2**: p. 86–92.
38. Mock, D.J., et al., *Leishmania induces survival, proliferation and elevated cellular dNTP levels in human monocytes promoting acceleration of HIV co-infection.* PLoS Pathog, 2012. **8**(4): p. e1002635.
39. WHO, *WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children. Geneva: World Health Organization.* 2006.
40. Ejara, E.D., et al., *Challenges in HIV and visceral Leishmania co-infection: future research directions.* Trop Med Int Health, 2010. **15**(10): p. 1266-7.
41. Yimer M, A.B., Mulu W, Zenebe Y, Bezabih B, *Proportion of visceral leishmaniasis and human immune deficiency virus co-infection among clinically confirmed visceral leishmaniasis patients at the endemic foci of the Amhara National Regional State, north-west Ethiopia.* Am J Biomed Life Sci, 2014. **2**: p. 1-7.

42. Brazil, M.o.H., *Manual de recomendações para diagnóstico, tratamento e acompanhamento de pacientes com a coinfecção leishmania-HIV.* 2011.
43. Burza, S., et al., *HIV and visceral leishmaniasis coinfection in Bihar, India: an underrecognized and underdiagnosed threat against elimination.* Clin Infect Dis, 2014. **59**(4): p. 552-5.
44. Stark, D., et al., *Post-kala-azar dermal leishmaniasis due to Leishmania infantum in a human immunodeficiency virus type 1-infected patient.* J Clin Microbiol, 2006. **44**(3): p. 1178-80.
45. WHO, *The post kala-azar dermal leishmaniasis (PKDL) atlas: a manual for health workers. Geneva: World Health Organization.* 2012.
46. Desjeux, P., et al., *Report of the Post Kala-Azar Dermal Leishmaniasis (PKDL) consortium meeting, New Delhi, India, 27–29 June 2012.* Parasites & Vectors, 2013. **6**(1): p. 196.
47. Reithinger, R., et al., *Cutaneous leishmaniasis.* Lancet Infect Dis, 2007. **7**(9): p. 581-96.
48. Kashif, M., et al., *Screening of Novel Inhibitors Against Leishmania donovani Calcium ion Channel to Fight Leishmaniasis.* Infect Disord Drug Targets, 2017. **17**(2): p. 120-129.
49. Miranda Gómez, O. and I. González Barea, *Leishmaniasis cutánea.* Revista Cubana de Medicina Militar, 2007. **36**: p. 0-0.
50. Bravo, F. and M.R. Sanchez, *New and re-emerging cutaneous infectious diseases in Latin America and other geographic areas.* Dermatol Clin, 2003. **21**(4): p. 655-68, viii.
51. Blanco, V.M., et al., *Clinical and epidemiologic profile of cutaneous leishmaniasis in Colombian children: considerations for local treatment.* Am J Trop Med Hyg, 2013. **89**(2): p. 359-64.
52. Alvar, J., et al., *Leishmaniasis worldwide and global estimates of its incidence.* PLoS One, 2012. **7**(5): p. e35671.
53. Kohn, G.C., *Encyclopedia of plague and pestilence : from ancient times to the present*. 3rd ed. 2008, New York: Facts On File. xiv, 529 p.
54. Ibrahim, M.E., *The epidemiology of visceral leishmaniasis in east Africa: hints and molecular revelations.* Trans R Soc Trop Med Hyg, 2002. **96 Suppl 1**: p. S25-9.
55. Franco, A.O., et al., *Predicting the distribution of canine leishmaniasis in western Europe based on environmental variables.* Parasitology, 2011. **138**(14): p. 1878-91.
56. WHO, *Regional strategic framework for elimination of kala-azar from the south-east Asia region (2005–2015). New Delhi: World Health Organization.* 2005.
57. Srivastava, P., et al., *Molecular and serological markers of Leishmania donovani infection in healthy individuals from endemic areas of Bihar, India.* Trop Med Int Health, 2013. **18**(5): p. 548-54.
58. Dos Santos Marques, L.H., et al., *Leishmania infantum: illness, transmission profile and risk factors for asymptomatic infection in an endemic metropolis in Brazil.* Parasitology, 2017. **144**(4): p. 546-556.
59. Shaw, J., *The leishmaniases--survival and expansion in a changing world. A mini-review.* Mem Inst Oswaldo Cruz, 2007. **102**(5): p. 541-7.
60. Thakur, C.P., *Treatment of kala-azar in India.* Natl Med J India, 1992. **5**(5): p. 203-5.
61. Sundar, S. and H.W. Murray, *Gamma interferon in the treatment of Kala-azar and other forms of Leishmaniasis.* J Assoc Physicians India, 1995. **43**(5): p. 348-50.
62. Arias, J.R., P.S. Monteiro, and F. Zicker, *The reemergence of visceral leishmaniasis in Brazil.* Emerg Infect Dis, 1996. **2**(2): p. 145-6.

63. Singh, O.P., et al., *Elimination of visceral leishmaniasis on the Indian subcontinent.* Lancet Infect Dis, 2016. **16**(12): p. e304-e309.
64. Wetzel, H. and D. Luger, *In vitro feeding in the rearing of tsetse flies (Glossina m. morsitans and G.p. palpalis, Diptera: Glossinidae).* Tropenmed Parasitol, 1978. **29**(2): p. 239-51.
65. NVBDCP, *National Road map for kala-azar elimination. New Delhi, India: Directorate of National Vector Borne Disease Control Programme.* 2014.
66. Addy, M. and A. Nandy, *Ten years of kala-azar in west Bengal, Part I. Did post-kala-azar dermal leishmaniasis initiate the outbreak in 24-Parganas?* Bull World Health Organ, 1992. **70**(3): p. 341-6.
67. Uranw, S., et al., *Post-kala-azar dermal leishmaniasis in Nepal: a retrospective cohort study (2000-2010).* PLoS Negl Trop Dis, 2011. **5**(12): p. e1433.
68. Maroli, M., et al., *Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern.* Med Vet Entomol, 2013. **27**(2): p. 123-47.
69. Ready, P.D., *Biology of phlebotomine sand flies as vectors of disease agents.* Annu Rev Entomol, 2013. **58**: p. 227-50.
70. Steverding, D., *The history of leishmaniasis.* Parasit Vectors, 2017. **10**(1): p. 82.
71. Bern, C., J.H. Maguire, and J. Alvar, *Complexities of assessing the disease burden attributable to leishmaniasis.* PLoS Negl Trop Dis, 2008. **2**(10): p. e313.
72. Killick-Kendrick, R., *The biology and control of phlebotomine sand flies.* Clin Dermatol, 1999. **17**(3): p. 279-89.
73. Herwaldt, B.L., *Leishmaniasis.* Lancet, 1999. **354**(9185): p. 1191-9.
74. Bates, P.A., *Leishmania sand fly interaction: progress and challenges.* Curr Opin Microbiol, 2008. **11**(4): p. 340-4.
75. Hommel, M., *Visceral leishmaniasis: biology of the parasite.* J Infect, 1999. **39**(2): p. 101-11.
76. Launois, P., et al., *Cytokines in parasitic diseases: the example of cutaneous leishmaniasis.* Int Rev Immunol, 1998. **17**(1-4): p. 157-80.
77. Solbach, W. and T. Laskay, *The host response to Leishmania infection.* Adv Immunol, 2000. **74**: p. 275-317.
78. Lehane, M.J., *Peritrophic matrix structure and function.* Annu Rev Entomol, 1997. **42**: p. 525-50.
79. Sunter, J. and K. Gull, *Shape, form, function and Leishmania pathogenicity: from textbook descriptions to biological understanding.* Open Biol, 2017. **7**(9).
80. Bates, P.A., *Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies.* Int J Parasitol, 2007. **37**(10): p. 1097-106.
81. Zilberstein, D. and M. Shapira, *The role of pH and temperature in the development of Leishmania parasites.* Annu Rev Microbiol, 1994. **48**: p. 449-70.
82. Gontijo, N.F., et al., *Lutzomyia longipalpis: pH in the gut, digestive glycosidases, and some speculations upon Leishmania development.* Exp Parasitol, 1998. **90**(3): p. 212-9.
83. Sundar, S., et al., *Failure of pentavalent antimony in visceral leishmaniasis in India: report from the center of the Indian epidemic.* Clin Infect Dis, 2000. **31**(4): p. 1104-7.
84. WHO, *Leishmaniasis. Geneva: World Health Organization.* 2010.
85. Dorlo, T.P., et al., *Optimal dosing of miltefosine in children and adults with visceral leishmaniasis.* Antimicrob Agents Chemother, 2012. **56**(7): p. 3864-72.

86. Kimutai, R., et al., *Safety and Effectiveness of Sodium Stibogluconate and Paromomycin Combination for the Treatment of Visceral Leishmaniasis in Eastern Africa: Results from a Pharmacovigilance Programme.* Clin Drug Investig, 2017. **37**(3): p. 259-272.
87. Balasegaram, M., et al., *Liposomal amphotericin B as a treatment for human leishmaniasis.* Expert Opin Emerg Drugs, 2012. **17**(4): p. 493-510.
88. Aggarwal V, B.B., Misra K, Tomar G, Burza S, *Decentralising AmBisome treatment for VL across endemic areas in India a novel nodal approach.* 2017.
89. Schlein, Y., *Sandfly diet and Leishmania.* Parasitol Today, 1986. **2**(6): p. 175-7.
90. Wäckers, F.L., *Suitability of (extra-)floral nectar, pollen, and honeydew as insect food sources,* in *Plant-Provided Food for Carnivorous Insects: A Protective Mutualism and its Applications,* F.L. Wäckers, J. Bruin, and P.C.J. van Rijn, Editors. 2005, Cambridge University Press: Cambridge. p. 17-74.
91. Jacobson, R.L., Y. Schlein, and C.L. Eisenberger, *The biological function of sand fly and Leishmania glycosidases.* Med Microbiol Immunol, 2001. **190**(1-2): p. 51-5.
92. Opperdoes, F., et al., *Leishmania, after the Genome.* 2008.
93. Ueno, N. and M.E. Wilson, *Receptor-mediated phagocytosis of Leishmania: implications for intracellular survival.* Trends Parasitol, 2012. **28**(8): p. 335-44.
94. Antoine, J.C., et al., *Parasitophorous vacuoles of Leishmania amazonensis-infected macrophages maintain an acidic pH.* Infect Immun, 1990. **58**(3): p. 779-87.
95. Antoine, J.C., et al., *Leishmania mexicana: a cytochemical and quantitative study of lysosomal enzymes in infected rat bone marrow-derived macrophages.* Exp Parasitol, 1987. **64**(3): p. 485-98.
96. Burchmore, R.J. and M.P. Barrett, *Life in vacuoles--nutrient acquisition by Leishmania amastigotes.* Int J Parasitol, 2001. **31**(12): p. 1311-20.
97. Landfear, S.M., *Nutrient transport and pathogenesis in selected parasitic protozoa.* Eukaryot Cell, 2011. **10**(4): p. 483-93.
98. Burchmore, R.J., et al., *Genetic characterization of glucose transporter function in Leishmania mexicana.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3901-6.
99. Hammond, D.J. and W.E. Gutteridge, *Purine and pyrimidine metabolism in the Trypanosomatidae.* Mol Biochem Parasitol, 1984. **13**(3): p. 243-61.
100. Pastakia, K.B. and D.M. Dwyer, *Identification and characterization of a ribose transport system in Leishmania donovani promastigotes.* Mol Biochem Parasitol, 1987. **26**(1-2): p. 175-81.
101. Ghosh, M. and T. Mukherjee, *Stage-specific development of a novel adenosine transporter in Leishmania donovani amastigotes.* Mol Biochem Parasitol, 2000. **108**(1): p. 93-9.
102. Carter, N.S., et al., *Cloning of a novel inosine-guanosine transporter gene from Leishmania donovani by functional rescue of a transport-deficient mutant.* J Biol Chem, 2000. **275**(27): p. 20935-41.
103. Vasudevan, G., et al., *Cloning of Leishmania nucleoside transporter genes by rescue of a transport-deficient mutant.* Proc Natl Acad Sci U S A, 1998. **95**(17): p. 9873-8.
104. Ortiz, D., et al., *Molecular genetic analysis of purine nucleobase transport in Leishmania major.* Mol Microbiol, 2007. **64**(5): p. 1228-43.
105. Ortiz, D., et al., *An acid-activated nucleobase transporter from Leishmania major.* J Biol Chem, 2009. **284**(24): p. 16164-9.
106. Krassner, S.M. and B. Flory, *Proline metabolism in Leishmania donovani promastigotes.* J Protozool, 1972. **19**(4): p. 682-5.

107. Hart, D.T. and G.H. Coombs, *Leishmania mexicana: energy metabolism of amastigotes and promastigotes.* Exp Parasitol, 1982. **54**(3): p. 397-409.
108. Basselin, M., G.H. Coombs, and M.P. Barrett, *Putrescine and spermidine transport in Leishmania.* Mol Biochem Parasitol, 2000. **109**(1): p. 37-46.
109. Drew, M.E., et al., *Functional expression of a myo-inositol/H+ symporter from Leishmania donovani.* Mol Cell Biol, 1995. **15**(10): p. 5508-15.
110. Kundig, C., et al., *Increased transport of pteridines compensates for mutations in the high affinity folate transporter and contributes to methotrexate resistance in the protozoan parasite Leishmania tarentolae.* EMBO J, 1999. **18**(9): p. 2342-51.
111. Saunders, E.C., et al., *Central carbon metabolism of Leishmania parasites.* Parasitology, 2010. **137**(9): p. 1303-13.
112. Naderer, T. and M.J. McConville, *The Leishmania-macrophage interaction: a metabolic perspective.* Cell Microbiol, 2008. **10**(2): p. 301-8.
113. Rodriguez-Contreras, D., et al., *Phenotypic characterization of a glucose transporter null mutant in Leishmania mexicana.* Mol Biochem Parasitol, 2007. **153**(1): p. 9-18.
114. Opperdoes, F.R. and J.P. Szikora, *In silico prediction of the glycosomal enzymes of Leishmania major and trypanosomes.* Mol Biochem Parasitol, 2006. **147**(2): p. 193-206.
115. Hellemond, J.J., B.M. Bakker, and A.G. Tielens, *Energy metabolism and its compartmentation in Trypanosoma brucei.* Adv Microb Physiol, 2005. **50**: p. 199-226.
116. Kovarova, J. and M.P. Barrett, *The Pentose Phosphate Pathway in Parasitic Trypanosomatids.* Trends Parasitol, 2016. **32**(8): p. 622-634.
117. Saunders, E.C., et al., *Isotopomer profiling of Leishmania mexicana promastigotes reveals important roles for succinate fermentation and aspartate uptake in tricarboxylic acid cycle (TCA) anaplerosis, glutamate synthesis, and growth.* J Biol Chem, 2011. **286**(31): p. 27706-17.
118. Guerra, D.G., et al., *The mitochondrial FAD-dependent glycerol-3-phosphate dehydrogenase of Trypanosomatidae and the glycosomal redox balance of insect stages of Trypanosoma brucei and Leishmania spp.* Mol Biochem Parasitol, 2006. **149**(2): p. 155-69.
119. Ralton, J.E., et al., *Evidence that intracellular beta1-2 mannan is a virulence factor in Leishmania parasites.* J Biol Chem, 2003. **278**(42): p. 40757-63.
120. Rodriguez-Contreras, D. and N. Hamilton, *Gluconeogenesis in Leishmania mexicana: contribution of glycerol kinase, phosphoenolpyruvate carboxykinase, and pyruvate phosphate dikinase.* J Biol Chem, 2014. **289**(47): p. 32989-3000.
121. Burchmore, R.J. and D.T. Hart, *Glucose transport in amastigotes and promastigotes of Leishmania mexicana mexicana.* Mol Biochem Parasitol, 1995. **74**(1): p. 77-86.
122. Rainey, P.M. and N.E. MacKenzie, *A carbon-13 nuclear magnetic resonance analysis of the products of glucose metabolism in Leishmania pifanoi amastigotes and promastigotes.* Mol Biochem Parasitol, 1991. **45**(2): p. 307-15.
123. Shier, W.T., et al., *(-)-BOTRYODIPLODIN, A UNIQUE RIBOSE-ANALOG TOXIN.* Toxin Reviews, 2007. **26**(4): p. 343-386.
124. Cameron, M.M., et al., *Sugar meal sources for the phlebotomine sandfly Lutzomyia longipalpis in Ceara State, Brazil.* Med Vet Entomol, 1995. **9**(3): p. 263-72.
125. Lardeux, B.R. and G.E. Mortimore, *Amino acid and hormonal control of macromolecular turnover in perfused rat liver. Evidence for selective autophagy.* J Biol Chem, 1987. **262**(30): p. 14514-9.

126. Steiger, R.F. and C.D. Black, *Simplified defined media for cultivating Leishmania donovani promastigotes.* Acta Trop, 1980. **37**(2): p. 195-8.
127. Cui, L., G.R. Rajasekariah, and S.K. Martin, *A nonspecific nucleoside hydrolase from Leishmania donovani: implications for purine salvage by the parasite.* Gene, 2001. **280**(1-2): p. 153-62.
128. Maugeri, D.A., et al., *Pentose phosphate metabolism in Leishmania mexicana.* Mol Biochem Parasitol, 2003. **130**(2): p. 117-25.
129. Burchmore, R.J. and S.M. Landfear, *Differential regulation of multiple glucose transporter genes in Leishmania mexicana.* J Biol Chem, 1998. **273**(44): p. 29118-26.
130. Naula, C.M., et al., *A glucose transporter can mediate ribose uptake: definition of residues that confer substrate specificity in a sugar transporter.* J Biol Chem, 2010. **285**(39): p. 29721-8.
131. Feng, X., et al., *Amplification of an alternate transporter gene suppresses the avirulent phenotype of glucose transporter null mutants in Leishmania mexicana.* Mol Microbiol, 2009. **71**(2): p. 369-81.
132. Marr, J.J., R.L. Berens, and D.J. Nelson, *Purine metabolism in Leishmania donovani and Leishmania braziliensis.* Biochim Biophys Acta, 1978. **544**(2): p. 360-71.
133. LaFon, S.W., et al., *Purine and pyrimidine salvage pathways in Leishmania donovani.* Biochem Pharmacol, 1982. **31**(2): p. 231-8.
134. Koszalka, G.W. and T.A. Krenitsky, *Nucleosidases from Leishmania donovani. Pyrimidine ribonucleosidase, purine ribonucleosidase, and a novel purine 2'-deoxyribonucleosidase.* J Biol Chem, 1979. **254**(17): p. 8185-93.
135. Alifano, P., et al., *Histidine biosynthetic pathway and genes: structure, regulation, and evolution.* Microbiol Rev, 1996. **60**(1): p. 44-69.
136. Riggs, J.W., et al., *Identification of the Plant Ribokinase and Discovery of a Role for Arabidopsis Ribokinase in Nucleoside Metabolism.* J Biol Chem, 2016. **291**(43): p. 22572-22582.
137. Hendrickson, N., T. Allen, and B. Ullman, *Molecular characterization of phosphoribosylpyrophosphate synthetase from Leishmania donovani.* Mol Biochem Parasitol, 1993. **59**(1): p. 15-27.
138. Eggleston, L.V. and H.A. Krebs, *Regulation of the pentose phosphate cycle.* Biochem J, 1974. **138**(3): p. 425-35.
139. Kovarova, J., et al., *Deletion of transketolase triggers a stringent metabolic response in promastigotes and loss of virulence in amastigotes of Leishmania mexicana.* PLoS Pathog, 2018. **14**(3): p. e1006953.
140. Conway, L.P. and J. Voglmeir, *Functional analysis of anomeric sugar kinases.* Carbohydr Res, 2016. **432**: p. 23-30.
141. Dyguda, E., B. Szefczyk, and W. Sokalski, *The Mechanism of Phosphoryl Transfer Reaction and the Role of Active Site Residues on the Basis of Ribokinase-Like Kinases.* International Journal of Molecular Sciences, 2004. **5**(4): p. 141.
142. Dyguda-Kazimierowicz, E., B. Szefczyk, and W. Sokalski, *The Mechanism of Phosphoryl Transfer Reaction and the Role of Active Site Residues on the Basis of Ribokinase-Like Kinases*. Vol. 5. 2004.
143. Bork, P., C. Sander, and A. Valencia, *Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases.* Protein Sci, 1993. **2**(1): p. 31-40.

144. Anderson, A. and R.A. Cooper, *The significance of ribokinase for ribose utilization by Escherichia coli.* Biochim Biophys Acta, 1969. **177**(1): p. 163-5.
145. Lopilato, J.E., et al., *D-ribose metabolism in Escherichia coli K-12: genetics, regulation, and transport.* J Bacteriol, 1984. **158**(2): p. 665-73.
146. Akpunarlieva, S., et al., *Integration of proteomics and metabolomics to elucidate metabolic adaptation in Leishmania.* J Proteomics, 2017. **155**: p. 85-98.
147. Kerkhoven, E.J., et al., *Handling uncertainty in dynamic models: the pentose phosphate pathway in Trypanosoma brucei.* PLoS Comput Biol, 2013. **9**(12): p. e1003371.
148. Sigrell, J.A., A.D. Cameron, and S.L. Mowbray, *Induced fit on sugar binding activates ribokinase.* J Mol Biol, 1999. **290**(5): p. 1009-18.
149. Li, J., et al., *Crystal structure of Sa239 reveals the structural basis for the activation of ribokinase by monovalent cations.* J Struct Biol, 2012. **177**(2): p. 578-82.
150. Paul, R., M.D. Patra, and U. Sen, *Crystal structure of apo and ligand bound vibrio cholerae ribokinase (Vc-RK): role of monovalent cation induced activation and structural flexibility in sugar phosphorylation.* Adv Exp Med Biol, 2015. **842**: p. 293-307.
151. Andersson, C.E. and S.L. Mowbray, *Activation of ribokinase by monovalent cations.* J Mol Biol, 2002. **315**(3): p. 409-19.
152. Ogbunude, P.O., N. Lamour, and M.P. Barrett, *Molecular cloning, expression and characterization of ribokinase of Leishmania major.* Acta Biochim Biophys Sin (Shanghai), 2007. **39**(6): p. 462-6.
153. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Methods Enzymol, 1997. **276**: p. 307-26.
154. Kabsch, W., *Integration, scaling, space-group assignment and post-refinement.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 133-44.
155. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 22-5.
156. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
157. Murshudov, G.N., et al., *REFMAC5 for the refinement of macromolecular crystal structures.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 355-67.
158. Vaguine, A.A., J. Richelle, and S.J. Wodak, *SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 1): p. 191-205.
159. Sigrell, J.A., et al., *Structure of Escherichia coli ribokinase in complex with ribose and dinucleotide determined to 1.8 A resolution: insights into a new family of kinase structures.* Structure, 1998. **6**(2): p. 183-93.
160. Parducci, R.E., et al., *Evidence for a catalytic Mg2+ ion and effect of phosphate on the activity of Escherichia coli phosphofructokinase-2: regulatory properties of a ribokinase family member.* Biochemistry, 2006. **45**(30): p. 9291-9.
161. Cabrera, R., J. Babul, and V. Guixe, *Ribokinase family evolution and the role of conserved residues at the active site of the PfkB subfamily representative, Pfk-2 from Escherichia coli.* Arch Biochem Biophys, 2010. **502**(1): p. 23-30.
162. Slabinski, L., et al., *XtalPred: a web server for prediction of protein crystallizability.* Bioinformatics, 2007. **23**(24): p. 3403-3405.
163. Gasteiger, E., et al., *Protein identification and analysis tools on the ExPASy server*, in *The proteomics protocols handbook*. 2005, Springer. p. 571-607.

164. Chuvikovsky, D.V., et al., *Ribokinase from E. coli: expression, purification, and substrate specificity.* Bioorg Med Chem, 2006. **14**(18): p. 6327-32.
165. Quiroga-Roger, D., J. Babul, and V. Guixe, *Role of monovalent and divalent metal cations in human ribokinase catalysis and regulation.* Biometals, 2015. **28**(2): p. 401-13.
166. Park, J., et al., *Identification and characterization of human ribokinase and comparison of its properties with E. coli ribokinase and human adenosine kinase.* FEBS Lett, 2007. **581**(17): p. 3211-6.
167. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases.* Biopolymers, 2008. **89**(5): p. 392-400.
168. Arnfors, L., et al., *Structure of Methanocaldococcus jannaschii nucleoside kinase: an archaeal member of the ribokinase family.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 9): p. 1085-97.
169. Carugo, O., *Buried chloride stereochemistry in the Protein Data Bank.* BMC Struct Biol, 2014. **14**: p. 19.
170. Gohara, D.W. and E. Di Cera, *Molecular Mechanisms of Enzyme Activation by Monovalent Cations.* J Biol Chem, 2016. **291**(40): p. 20840-20848.
171. Johal, A.R., et al., *Sequence-dependent effects of cryoprotectants on the active sites of the human ABO(H) blood group A and B glycosyltransferases.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 3): p. 268-76.
172. Schumacher, M.A., et al., *Crystal structures of Toxoplasma gondii adenosine kinase reveal a novel catalytic mechanism and prodrug binding.* J Mol Biol, 2000. **298**(5): p. 875-93.
173. Alfaro, J.A., et al., *ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes.* J Biol Chem, 2008. **283**(15): p. 10097-108.
174. Espina, G., et al., *A novel beta-xylosidase structure from Geobacillus thermoglucosidasius: the first crystal structure of a glycoside hydrolase family GH52 enzyme reveals unpredicted similarity to other glycoside hydrolase folds.* Acta Crystallogr D Biol Crystallogr, 2014. **70**(Pt 5): p. 1366-74.
175. Lee, J. and M. Paetzel, *Structure of the catalytic domain of glucoamylase from Aspergillus niger.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2011. **67**(Pt 2): p. 188-92.
176. Nogly, P., et al., *High-resolution structure of an atypical alpha-phosphoglucomutase related to eukaryotic phosphomannomutases.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 10): p. 2008-16.

8. Publications

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Ribokinase from *Leishmania donovani*: purification, characterization and X-ray crystallographic analysis

Santhosh Gatreddi, Sayanna Are and Insaf Ahmed Qureshi*

Department of Biotechnology and Bioinformatics, School of Life Sciences, University of Hyderabad, Professor C. R. Rao Road, Hyderabad 500 046, India. *Correspondence e-mail: insaf@uohyd.ac.in

Received 23 October 2017
Accepted 2 January 2018

Edited by P. Dunten, Stanford Synchrotron Radiation Lightsource, USA

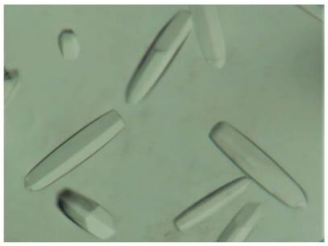
Keywords: *Leishmania donovani*; D-ribose metabolism; ribokinase; crystallization; thermal stability.

Leishmania is an auxotrophic protozoan parasite which acquires D-ribose by transporting it from the host cell and also by the hydrolysis of nucleosides. The enzyme ribokinase (RK) catalyzes the first step of ribose metabolism by phosphorylating D-ribose using ATP to produce D-ribose-5-phosphate. To understand its structure and function, the gene encoding RK from *L. donovani* was cloned, expressed and purified using affinity and size-exclusion chromatography. Circular-dichroism spectroscopy of the purified protein showed comparatively more α-helix in the secondary-structure content, and thermal unfolding revealed the T_m to be 317.2 K. Kinetic parameters were obtained by functional characterization of *L. donovani* RK, and the K_m values for ribose and ATP were found to be 296 $\pm$ 36 and 116 $\pm$ 9.0 μM, respectively. Crystals obtained by the hanging-drop vapour-diffusion method diffracted to 1.95 Å resolution and belonged to the hexagonal space group $P6_1$, with unit-cell parameters $a = b = 100.25$, $c = 126.77$ Å. Analysis of the crystal content indicated the presence of two protomers in the asymmetric unit, with a Matthews coefficient (V_M) of 2.45 Å^3 Da^{-1} and 49.8% solvent content. Further study revealed that human counterpart of this protein could be used as a template to determine the first three-dimensional structure of the RK from trypanosomatid parasites.

1. Introduction

D-Ribose is an abundant pentose sugar in biological systems that can be used as a source of energy as well as as a constituent of RNA, DNA and nucleotide cofactors. Phosphorylation of ribose mediated by ribokinase (RK; EC 2.7.1.15) is the first step in its entrance into metabolic pathways and also into the biosynthesis of nucleotides, tryptophan and histidine (Anderson & Cooper, 1969; Lopilato *et al.*, 1984). Analysis of ribokinase gene sequences from both prokaryotes and eukaryotes identified that they belong to the PfkB family of carbohydrate kinases, also known as the RK family (Bork *et al.*, 1993; Park & Gupta, 2008).

Ribokinase catalyzes the transfer of the γ-phosphate of ATP to O5′ of D-ribose to produce ADP and D-ribose-5-phosphate. The catalytic activity of ribokinase is regulated by monovalent and divalent cations. Biochemical studies of ribokinases from *Escherichia coli*, *Arabidopsis thaliana* and human revealed that divalent metal cations are critical for catalytic activity, preferably Mg^{2+} or Mn^{2+} (Chuvikovsky *et al.*, 2006; Quiroga-Roger *et al.*, 2015; Riggs *et al.*, 2016). Metal cations other than Mg^{2+} and Mn^{2+}, such as Co^{2+}, Ca^{2+}, Cu^{2+}, Cd^{2+}, Ni^{2+} and Zn^{2+}, also support the activity of ribokinase (Chuvikovsky *et al.*, 2006; Quiroga-Roger *et al.*, 2015; Schimmel *et al.*, 1974). Divalent metal cations regulate activity by forming a metal–nucleotide complex and also by binding to the conserved N*XX*E motif present near the active site of



ribokinase (Quiroga-Roger *et al.*, 2015). On the other hand, monovalent cations, preferably K^+, Cs^+ and NH_4^+, modulate catalysis by activating ribokinase (Andersson & Mowbray, 2002; Chuvikovsky *et al.*, 2006). The mechanism of activation by monovalent cations has been demonstrated by the binding of a cesium ion between two loops near to the active site in the crystal structures of ribokinase from *E. coli* and *Vibrio cholerae* (Andersson & Mowbray, 2002; Paul *et al.*, 2015). The two structural changes that were associated with the binding of monovalent cation to the enzyme comprised a change in the conformation of the large ATP loop, increasing the affinity for ATP, and the formation of an anion hole to stabilize the transition state (Andersson & Mowbray, 2002; Li *et al.*, 2012).

Visceral leishmaniasis is a clinical syndrome caused by the protozoan parasite *Leishmania donovani*, which belongs to the Trypanosomatidae family. *Leishmania* exists in two morphological forms: a motile promastigote that resides in the midgut of phlebotomine sandflies and an intracellular amastigote that is present in the parasitophorous vacuole of the mammalian macrophage (Sunter & Gull, 2017). For rapid growth and survival, the parasite acquires most of its vital metabolites from the sandfly midgut and the parasitophorous vacuoles of the macrophage during the promastigote and amastigote stages of its lifecycle, respectively (Chang & Dwyer, 1978; Basselin *et al.*, 2000; Burchmore & Hart, 1995; Drew *et al.*, 1995; Ghosh & Mukherjee, 2000; Kündig *et al.*, 1999; Mazareb *et al.*, 1999; Zilberstein & Gepstein, 1993). Among these metabolites, D-ribose accumulates in the promastigotes of *L. donovani* by a carrier-mediated process and also through the degradation of nucleosides by non-specific nucleoside hydrolase (Cui *et al.*, 2001; Pastakia & Dwyer, 1987; Naula *et al.*, 2010). Experiments using radio-labelled D-ribose with *L. mexicana* have identified the importance of ribokinase in the incorporation of free ribose into nucleic acids (Maugeri *et al.*, 2003). Ribokinase plays a key role by phosphorylating the ribose available in the cytosol to enable its entry into *Leishmania* metabolism. To understand the catalytic function using a structure-based approach, the ribokinase gene from *L. donovani* (*Ld*RK) was cloned, purified, characterized and crystallographic analysis was carried out.

2. Materials and methods

2.1. Cloning, expression and purification

The full-length ORF for ribokinase was amplified from the genomic DNA of *L. donovani* by polymerase chain reaction using the oligonucleotide primers in Table 1. The amplified PCR product was purified using a gel-extraction kit (Thermo Scientific) and ligated into pJET1.2 cloning vector (Thermo Scientific). The insert was released from the pJET1.2 vector by digestion with the NdeI and XhoI restriction enzymes and was then subcloned into pET-28a(+) (Novagen) expression vector, which adds a hexahistidine tag to the N-terminus of the target protein. The plasmids isolated from the positive clones were verified by restriction digestion followed by sequencing prior to transformation into *E. coli* BL21 (DE3) cells. For over-expression, a single colony was inoculated into 30 ml Luria–Bertani (LB) medium containing 50 µg ml^{-1} kanamycin and incubated overnight at 310 K. An overnight-grown culture (20 ml) was diluted to 1% with LB broth, incubated at 310 K to an OD_{600} of 0.6 and then induced with 0.05 m*M* IPTG for 20 h at 291 K. Induced cells were pelleted for 10 min at 6300*g* and resuspended in 50 ml lysis buffer (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 40 m*M* imidazole, 2 m*M* β-mercapto-ethanol, 0.25 mg ml^{-1} lysozyme, 0.1 m*M* phenylmethyl-sulfonyl fluoride) followed by lysis using sonication. The cell lysate was centrifuged at 21 000*g* for 30 min, and the supernatant was filtered using a 0.2 µm filter (Millipore) and loaded onto a HisTrap HP 5 ml column (GE Healthcare) equilibrated with buffer *A* consisting of 50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl. Unbound bacterial proteins were eluted with a step gradient of 10, 15, 20 and 30% buffer *B* (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 500 m*M* imidazole) and the desired protein was eluted from the column using 80% buffer *B*. Eluted fractions with high purity were pooled and loaded onto a HiLoad 16/60 Superdex 200 pg column (GE Healthcare) equilibrated with 50 m*M* Tris–HCl pH 8.0, 125 m*M* NaCl. The fractions with the highest purity were pooled and then concentrated using Amicon Ultra (10 kDa cutoff) to 8 mg ml^{-1} for crystallization. All purification steps were performed at 277 K and protein concentrations were determined at 280 nm with a NanoDrop 2000 (Thermo Scientific) using a molar extinction coefficient (ε) of 16 390 M^{-1} cm^{-1} and the molecular weight.

Table 1
Macromolecule-production information.

Source organism	*L. donovani*
DNA source	Genomic DNA
Cloning vector	pJET1.2
Expression vector	pET-28a(+)
Forward primer†	5′-AATCATATGCACCGTGCGCGGAACGTT-3′
Reverse primer‡	5′-AATCTCGAGCTACATGACACCAGCCGGCA-3′
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced§	MGSSHHHHHHSSGLVPRGSHMHRARNVRSHTGEY APDILVVGSCFLDYVGYVDHMPQVGETMHSVS FHKGFGGKGANQAVAAGRLGAKVAMVSMVGTD GDGSDYIKELERNGVDTAYMFRTGKSSTGLAM ILVDTKSSNNEIVICPNATNHFTPELLRAQTN NYERILHTGLKYLICQNEIPLPTTLDTIKEAH SRGVYTVFNSAPAPKPAEVEQIKPFLPYVSLF CPNEVEATLITGVKVTDTESAFSAIKALQQLG VRDVVITLGAAGFVLSENGAEPVHVTGKHVKA VDTTGAGDCFVGSMVYFMSRGRNLLEACKRAN ECAAISVTRKGTQLSYPHPSELPAGVM

† The NdeI site is underlined. ‡ The XhoI site is underlined. § Non-native residues originating from the vector are underlined.

2.2. Circular dichroism

The circular-dichroism (CD) spectrum was measured using 7.5 m*M* Tris–HCl pH 8.5 containing 3.2 µ*M* *Ld*RK with a J-1500 spectropolarimeter (Jasco) at 298 K from 260 to 190 nm with a scanning speed of 20 nm min^{-1}. The path length of the sample cell was 2 mm, while the data interval and bandwidth were both 1 nm. Two consecutive spectra were recorded and

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well plate
Temperature (K)	296
Protein concentration (mg ml^{-1})	8.0
Buffer composition of protein solution	50 m*M* Tris–HCl pH 8.0, 125 m*M* NaCl
Composition of reservoir solution	0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl, 4.5%(*v*/*v*) glycerol
Volume and ratio of drop	4 µl (1:1 protein:reservoir solution)
Volume of reservoir (µl)	500

averaged, and the data were plotted after subtracting the buffer blank in units of mdeg. Secondary-structure content was evaluated using the *CDSSTR* algorithm with reference data set No. 4 from the *DichroWeb* server (Whitmore & Wallace, 2008). Thermal stability was monitored in 10 m*M* sodium phosphate buffer pH 7.4 by measuring the ellipticity at 222 nm from 293 to 343 K with a temperature gradient of 1 K min^{-1}. The melting temperature (T_m) was calculated by plotting the first derivative of the thermal denaturation curve. All spectra were smoothed before plotting.

2.3. Enzyme kinetics

Enzyme assays were performed in triplicate at 298 K in a 96-well plate using a microplate reader (Molecular Devices) with a path length of 0.596 cm. The activity of ribokinase was measured using the method reported previously (Ogbunude *et al.*, 2007) with some modifications, by coupling the production of ADP to the oxidation of NADH and measuring the change in absorbance at 340 nm. The reaction mixture consisted of 30 m*M* Tris–HCl pH 8.5, 100 m*M* KCl, 3 m*M* $MgCl_2$, 0.3 m*M* phosphoenolpyruvate, 0.2 m*M* NADH, with 0.4 U pyruvate kinase and 1.2 U lactate dehydrogenase as coupling enzymes. For determination of the initial reaction rates, the concentration of ATP was fixed at 0.8 m*M* and ribose was varied from 0.01 to 4 m*M*; when the concentration of ribose was fixed at 4 m*M*, that of ATP was varied from 0.005 to 0.8 m*M*. The reactions were started by the addition of 0.2 µ*M* enzyme and the absorbance was recorded every 10 s for 30 min. The kinetic data were fitted to the Michaelis–Menten equation through nonlinear regression using *GraphPad Prism*.

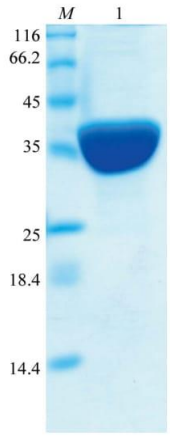


Figure 1
12% SDS–PAGE of purified recombinant *Ld*RK after gel-filtration chromatography. Lane *M*, molecular-weight markers (labelled in kDa); lane 1, pooled purified fraction from gel-filtration chromatography.

Table 3
Data collection and processing.

Values in parentheses are for the highest resolution shell.

Diffraction source	PX-BL21, Indus-2, RRCAT
Wavelength (Å)	0.9794
Temperature (K)	100
Detector	MAR 225 CCD
Crystal-to-detector distance (mm)	180.00
Rotation range per image (°)	1
Total rotation range (°)	118
Exposure time per image (s)	20
Space group	$P6_1$
a, *b*, *c* (Å)	100.25, 100.25, 126.77
α, β, γ (°)	90.0, 90.0, 120.0
Resolution range (Å)	46.61–1.95 (2.0–1.95)
Total No. of reflections	395500 (27813)
No. of unique reflections	52575 (3698)
Completeness (%)	100 (99.9)
Multiplicity	7.5 (7.5)
Mean $I/\sigma(I)$	14.9 (3.4)
R_{merge} (all *I*+ and *I*−)	0.101 (0.604)
$CC_{1/2}$	0.998 (0.876)
Overall *B* factor from Wilson plot (Å^2)	16.7

2.4. Crystallization

Crystallization screening was initially carried out by the sitting-drop vapour-diffusion method at 296 K using a Mosquito protein crystallization robot (TTP Labtech) at IICT, Hyderabad in MRC 96-well plates by mixing 0.15 µl *Ld*RK solution with 0.15 µl precipitant solution. Commercially available crystallization kits from Hampton Research (Index, PEG/Ion, SaltRx, Crystal Screen and Crystal Screen 2) and Molecular Dimensions (Structure Screen 1, Structure Screen 2, MORPHEUS, MIDAS, PGALM, PACT *premier* and JCSG *plus*) were used to screen for suitable conditions. Hexagonal rod-shaped crystals appeared after 1 d in a condition consisting of 0.1 *M* citric acid pH 3.5, 3.0 *M* NaCl. Further optimization was performed using the hanging-drop vapour-diffusion method by varying the concentration of protein, the pH of the buffer and the ionic strength of NaCl in 24-well plates with a reservoir volume of 500 µl. Crystals suitable for diffraction experiments were obtained after 21 d using 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl, 4.5%(*v*/*v*) glycerol (Table 2). Simultaneously, to confirm the presence of RK, crystals were washed with well solution, dissolved in sample-loading buffer and electrophoresed on 12% SDS–PAGE gels followed by silver staining.

2.5. Data collection and processing

A single crystal mounted on a loop was flash-cooled in liquid nitrogen with mother liquor as the cryoprotectant. Diffraction data were collected under cryogenic condition (100 K) on the protein crystallography beamline (PX-BL21)

at the Indus-2 synchrotron, Indore, India using a MAR 225 CCD detector. The collected diffraction data were processed using *XDS* (Kabsch, 2010) and scaled with *AIMLESS* from the *CCP*4 program suite (Evans & Murshudov, 2013). Data-collection and processing statistics are given in Table 3.

3. Results and discussion

Recombinant *Ld*RK was expressed in *E. coli* BL21 (DE3) and then purified using Ni^{2+}-affinity and gel-filtration chromatography, with a final yield of 15 mg per litre of culture medium. Analysis of the pooled purified fraction by SDS–PAGE indicated that the protein was >95% pure, with an approximate molecular weight of 38 kDa (Fig. 1), which corresponds to the theoretical molecular mass (37.5 kDa) predicted by *ProtParam* (Gasteiger *et al.*, 2005). Before proceeding to the crystallization of *Ld*RK, the secondary-structure content of the purified protein was analyzed by far-UV CD spectroscopy. The spectrum specifies α-helix as the dominant secondary structure in *Ld*RK, owing to the presence of the characteristic double minimum near 222 and 208 nm and a maximum around 192 nm. Analysis of the data using the *DichroWeb* server predicted the secondary-structure content to be as follows: 27% α-helix, 23% β-sheet, 22% turns and 28% unordered regions (Fig. 2*a*). The presence of a proportionately greater amount of α-helix in *Ld*RK correlates with the crystal structures of ribokinase reported from *V. cholerae* (Paul *et al.*, 2015), *Staphylococcus aureus* (Li *et al.*, 2012) and *E. coli* (Sigrell *et al.*, 1999). To probe the thermal stability of ribokinase, the change in the ellipticity was measured at 222 nm as a function of temperature from 293 to 343 K. The thermal melting (T_m) curve of ribokinase indicates a rapid increase in the ellipticity at 313 K, which begins to level off at 323 K (Fig. 2*b*). The T_m of the protein was calculated to be 317.2 K by plotting the first derivative of the thermal melting curve (Fig. 2*c*). The kinetic parameters (K_m, k_{cat} and k_{cat}/K_m) were obtained by plotting the initial velocities *versus* the respective substrate concentrations (Fig. 3), and the K_m values indicated that *Ld*RK has a higher affinity for ATP than ribose (Table 4). *Ld*RK binds to ribose with a K_m of 296 μM, which corroborates with the binding affinities (K_m) determined for RK from *L. major* (300 μM) and human (279 μM). The affinity of *Ld*RK towards ATP was observed to be similar (116 μM) to those of RK from *L. major* (200 μM; Ogbunude *et al.*, 2007) and human (70 μM; Chuvikovsky *et al.*, 2006).

Table 4
Kinetic parameters of *Ld*RK.

The parameters were obtained by performing the assays in triplicate.

Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} s^{-1})
Ribose	2.28 ± 0.08	296 ± 36	(7.70 ± 0.98) × 10^{-3}
ATP	2.15 ± 0.05	116 ± 9	(18.50 ± 1.50) × 10^{-3}

Although more than 950 conditions were screened to identify the preliminary crystallization conditions, crystals of *Ld*RK were only obtained with Index condition A7 consisting of 0.1 *M* citric acid pH 3.5, 3.0 *M* NaCl. Since the crystals that were obtained by screening were not suitable for diffraction studies, optimization was performed by varying the protein concentration, the pH and the NaCl concentration of the precipitant solution. After refining the initial crystallization conditions, long hexagonal rod-shaped crystals were obtained using 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl, 4.5% glycerol and reached their maximum size after three weeks (Fig. 4*a*). The presence of a protein band at the corresponding molecular weight (~38 kDa) on a silver-stained gel confirmed the

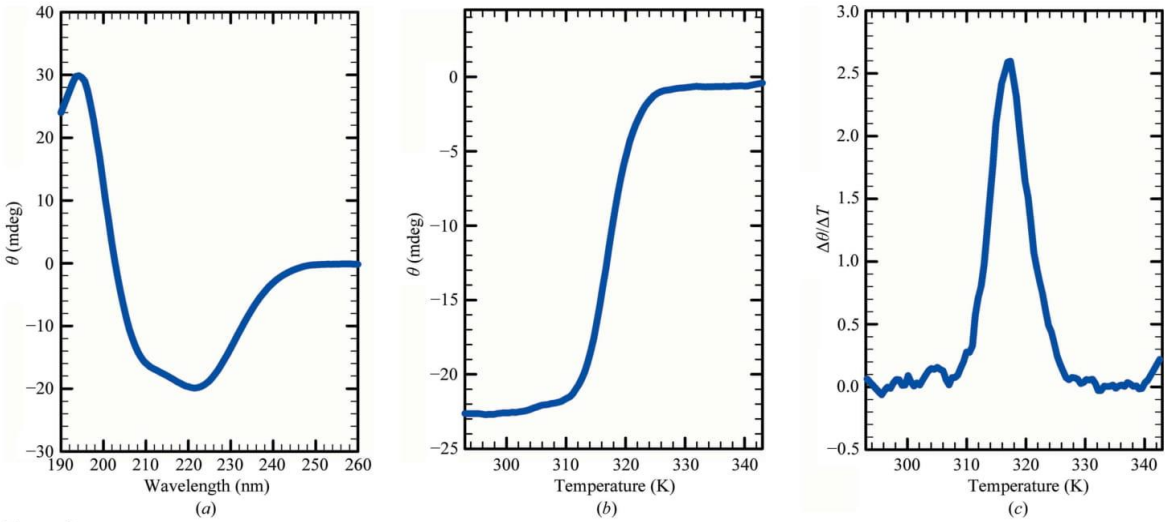


Figure 2
(*a*) Far-UV circular dichroism spectrum of *Ld*RK measured at 298 K in 7.5 m*M* Tris–HCl pH 8.5. (*b*) CD thermal denaturation curve of *Ld*RK obtained by measuring the ellipticity at 222 nm in 1.0 K intervals from 293 to 343 K. (*c*) The first-derivative curve of thermal melting with a peak value (T_m) of 317.2 K.

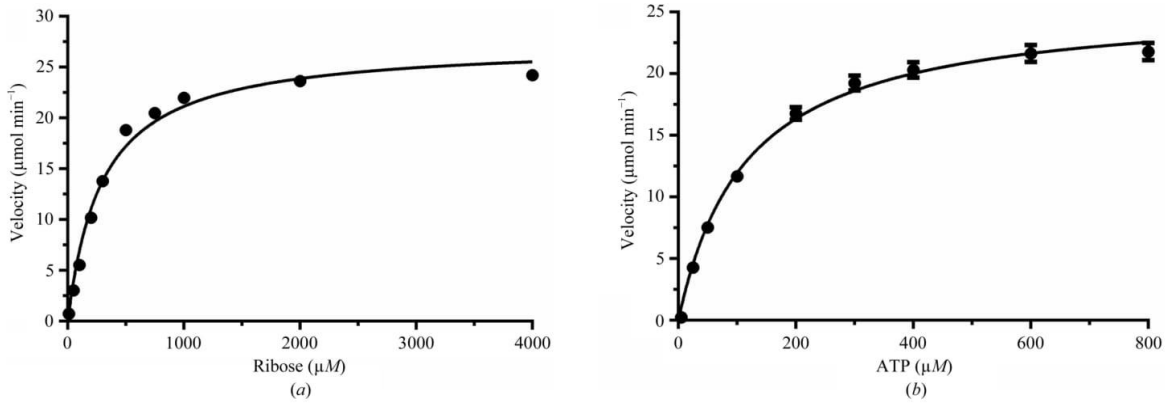


Figure 3
Kinetic characterization of *Ld*RK. (*a*) Plot of initial velocities against increasing concentrations of ribose. (*b*) Plot of initial velocities against increasing concentrations of ATP. Each data point represents the average from assays performed in triplicate; the standard error is shown as an error bar.

presence of *Ld*RK in the crystals (Fig. 4*b*). The addition of glycerol tremendously improved the crystal size, and these crystals diffracted to 1.95 Å resolution on the PX-BL21 synchrotron beamline (Fig. 5). The crystal lattice belonged to space group $P6_1$, with unit-cell parameters $a = b = 100.25$, $c = 126.77$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$ (Table 3). Two molecules of *Ld*RK were assumed to be present in the asymmetric unit, with a solvent content of 49.8% and a calculated Matthews coefficient (Matthews, 1968) of 2.45 Å^3 Da^{-1}. Using a *BLASTP* search of the PDB, ribokinase from *Homo sapiens* (PDB entry 2fv7; Structural Genomics Consortium, unpublished work) was identified as the closest homologous structure, with 42% identity to *Ld*RK. Molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using chain *A* of human

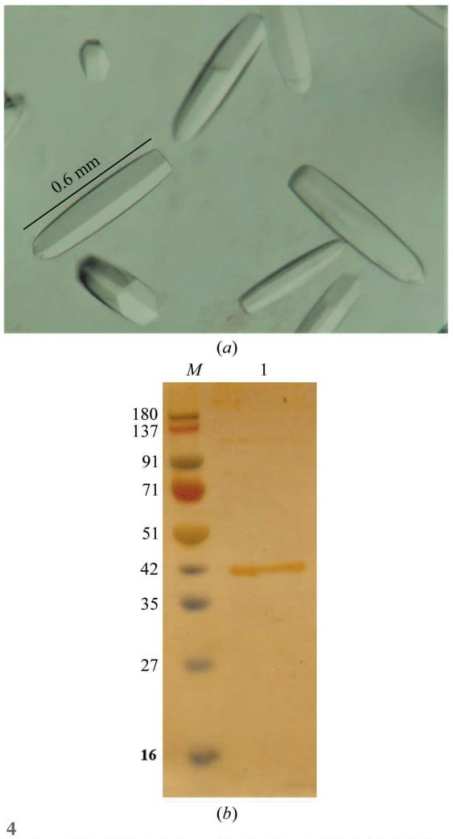


Figure 4
Crystallization of *Ld*RK. (*a*) Crystals obtained with 0.1 *M* citric acid pH 4.3, 3.4 *M* NaCl, 4.5%(*v*/*v*) glycerol using the hanging-drop vapour-diffusion method. (*b*) Silver-stained SDS–PAGE of the crystals. Lane *M*, molecular-weight markers (labelled in kDa); lane 1, crystals of *Ld*RK.

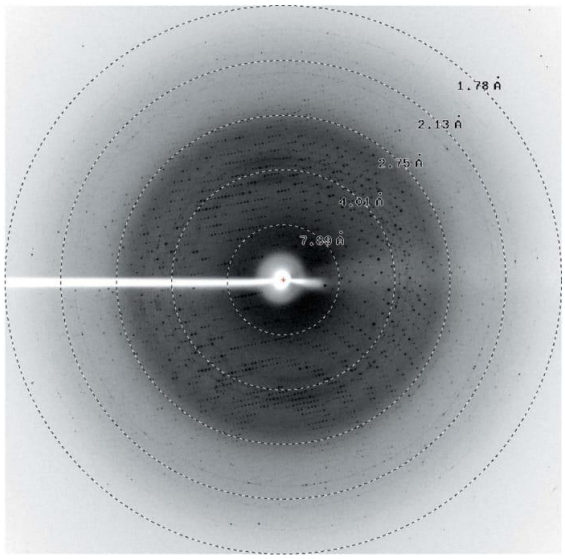


Figure 5
Diffraction image of *Ld*RK at 1.95 Å resolution obtained using a MAR 225 CCD detector.

RK (PDB entry 2fv7) as the search model suggested the presence of two protomers in the asymmetric unit. The solution obtained in space group $P6_1$ was confirmed by comparing the results of molecular replacement in space groups $P6_1$ (correlation coefficient of 0.431 and R factor of 59.5%) and $P6_5$ (correlation coefficient of 0.364 and R factor of 62.2%). After one round of initial restrained refinement, the R factor decreased to 42.1% in space group $P6_1$. The structure solution and refinement of this protein structure, which are in progress, will help us to understand the detailed catalytic mechanism of this enzyme in the ribose metabolism of leishmanial parasites.

Acknowledgements

We thank Radheshyam Maurya and Dileep Vasudevan for providing the genomic DNA of *L. donovani* and for critical reading of the manuscript, respectively. We are grateful to Anthony Addlagatta for letting us use the Protein Crystallization Robotic facility at IICT, Hyderabad, India. We are also grateful to the staff of the PX-BL21 beamline (DAE) at Indus-2, RRCAT, Indore, India for their help in data collection.

Funding information

We would like to thank the Science and Engineering Research Board, India (SERB; project No. SR/FT/LS-112/2011) and UPE-II, University of Hyderabad for financial assistance. SG and SA would like to acknowledge UGC and CSIR, respectively, for the award of fellowships.

References

Anderson, A. & Cooper, R. A. (1969). *Biochim. Biophys. Acta*, **177**, 163–165.
Andersson, C. E. & Mowbray, S. L. (2002). *J. Mol. Biol.* **315**, 409–419.
Basselin, M., Coombs, G. H. & Barrett, M. P. (2000). *Mol. Biochem. Parasitol.* **109**, 37–46.
Bork, P., Sander, C. & Valencia, A. (1993). *Protein Sci.* **2**, 31–40.
Burchmore, R. J. & Hart, D. T. (1995). *Mol. Biochem. Parasitol.* **74**, 77–86.
Chang, K. P. & Dwyer, D. M. (1978). *J. Exp. Med.* **147**, 515–530.
Chuvikovsky, D. V., Esipov, R. S., Skoblov, Y. S., Chupova, L. A., Muravyova, T. I., Miroshnikov, A. I., Lapinjoki, S. & Mikhailopulo, I. A. (2006). *Bioorg. Med. Chem.* **14**, 6327–6332.
Cui, L., Rajasekariah, G. R. & Martin, S. K. (2001). *Gene*, **280**, 153–162.
Drew, M. E., Langford, C. K., Klamo, E. M., Russell, D. G., Kavanaugh, M. P. & Landfear, S. M. (1995). *Mol. Cell. Biol.* **15**, 5508–5515.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa, NJ: Humana Press.
Ghosh, M. & Mukherjee, T. (2000). *Mol. Biochem. Parasitol.* **108**, 93–99.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kündig, C., Haimeur, A., Légaré, D., Papadopoulou, B. & Ouellette, M. (1999). *EMBO J.* **18**, 2342–2351.
Li, J., Wang, C., Wu, Y., Wu, M., Wang, L., Wang, Y. & Zang, J. (2012). *J. Struct. Biol.* **177**, 578–582.
Lopilato, J. E., Garwin, J. L., Emr, S. D., Silhavy, T. J. & Beckwith, J. R. (1984). *J. Bacteriol.* **158**, 665–673.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Maugeri, D. A., Cazzulo, J. J., Burchmore, R. J., Barrett, M. P. & Ogbunude, P. O. (2003). *Mol. Biochem. Parasitol.* **130**, 117–125.
Mazareb, S., Fu, Z. Y. & Zilberstein, D. (1999). *Exp. Parasitol.* **91**, 341–348.
Naula, C. M., Logan, F. M., Wong, P. E., Barrett, M. P. & Burchmore, R. J. (2010). *J. Biol. Chem.* **285**, 29721–29728.
Ogbunude, P. O., Lamour, N. & Barrett, M. P. (2007). *Acta Biochim. Biophys. Sin.* **39**, 462–466.
Park, J. & Gupta, R. S. (2008). *Cell. Mol. Life Sci.* **65**, 2875–2896.
Pastakia, K. B. & Dwyer, D. M. (1987). *Mol. Biochem. Parasitol.* **26**, 175–181.
Paul, R., Patra, M. D. & Sen, U. (2015). *Adv. Exp. Med. Biol.* **842**, 293–307.
Quiroga-Roger, D., Babul, J. & Guixé, V. (2015). *Biometals*, **28**, 401–413.
Riggs, J. W., Rockwell, N. C., Cavales, P. C. & Callis, J. (2016). *J. Biol. Chem.* **291**, 22572–22582.
Schimmel, S. D., Hoffee, P. & Horecker, B. L. (1974). *Arch. Biochem. Biophys.* **164**, 560–570.
Sigrell, J. A., Cameron, A. D. & Mowbray, S. L. (1999). *J. Mol. Biol.* **290**, 1009–1018.
Sunter, J. & Gull, K. (2017). *Open Biol.* **7**, 170165.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Whitmore, L. & Wallace, B. A. (2008). *Biopolymers*, **89**, 392–400.
Zilberstein, D. & Gepstein, A. (1993). *Mol. Biochem. Parasitol.* **61**, 197–205.

Understanding the catalytic mechanism of ribokinase from Leishmania donovani using structural and functional approach

by Santhosh Gatreddi

Submission date: 30-Nov-2018 12:27PM (UTC+0530)
Submission ID: 1047700905
File name: Thesis_29_11_18.docx (18.23M)
Word count: 15598
Character count: 84505

Understanding the catalytic mechanism of ribokinase from Leishmania donovani using structural and functional approach

ORIGINALITY REPORT

16%	13%	15%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journals.iucr.org Internet Source	10%
2	Sakib Burza, Simon L Croft, Marleen Boelaert. "Leishmaniasis", The Lancet, 2018 Publication	1%
3	David Pace. "Leishmaniasis", Journal of Infection, 2014 Publication	<1%
4	Chinmayee Mohapatra, Manas Kumar Jagdev, Dileep Vasudevan. "Crystal structures reveal N-terminal Domain of Arabidopsis thaliana ClpD to be highly divergent from that of ClpC1", Scientific Reports, 2017 Publication	<1%
5	docplayer.net Internet Source	<1%
6	"Kala Azar in South Asia", Springer Nature, 2016 Publication	<1%

University of Hyderabad
Hyderabad- 500 046, India

PLAGIARISM FREE CERTIFICATE

This is to certify that the similarity index of this thesis as checked by the Library of University of Hyderabad is 16%. Out of this, 10% similarity has been found to be identified from the candidate's own publication (Santhosh Gatreddi) which forms the adequate part of the thesis. The details of student's publication are as follows:

1. Santhosh Gatreddi, Sayanna Are and Insaf Ahmed Qureshi. "Ribokinase from *Leishmania donovani*: purification, characterization and X-ray crystallographic analysis." Acta Crystallographica Section F: Structural Biology Communications, 99:104, 2018.	**% Similarity** **10**

About 6% similarity was identified from external sources in the present thesis which is according to prescribed regulations of the University. All the publications related to this thesis have been appended at the end of the thesis. Hence the present thesis may be considered to be plagiarism free.

Dr. Insaf Ahmed Qureshi
Supervisor